CLONING, EXPRESSION AND PURIFICATION OF RECOMBINANT ELONGATION FACTOR Ts FROM HYPERTHERMOPHILIC BACTERIA *GEOBACILLUS ANATOLICUS*

by

Toros Şahin

B.S. in Molecular Biology and Genetics, Boğaziçi University, 2003

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Molecular Biology and Genetics
Boğaziçi University
2006

Aileme

ve

birtanecik aşkım Begüm'e

ACKNOWLEDGEMENTS

This work is supported by a grant from Boğaziçi University Research Fond (BAP project no: 03S102). I would like to thank my thesis advisor Prof. Neş'e Bilgin for her endless support, guidance and perhaps the most important, passion for science which passed on me during this study in her lab.

I am grateful to Prof. Ahmet Koman and Prof. Ayşe Özer for spending their time on evaluation of my work as my thesis jury.

I am also grateful to my lab mates Tunç Öztunç, Memed Doğan and Melike Çağlayan for their help thoughout this work.

Barış Akalın was great friend and great scientist throughout this work, he always gave me a hand with his valuable discussions when i was stuck at dead-ends. Thank you Barış.

I am also grateful to my family. Without their support, this work couldn't have been done.

My final thanks go to my endless passion Begüm Erdil, who was always understanding, patient and helpful during my work. She was always with me, always caring for me. I know she will be with me forever, thank you Begüm.

ABSTRACT

CLONING, EXPRESSION AND PURIFICATION OF RECOMBINANT ELONGATION FACTOR TS FROM HYPERTHERMOPHILIC BACTERIA *GEOBACILLUS ANATOLICUS*

Elongation factor EF-Ts from *Geobacillus anatolicus* was cloned, expressed and purifed as a His-tagged recombinant protein in *Escherichia coli*. The first step in this study was the determination of the nucleic acid sequence of the *tsf* gene of *Geobacillus anatolicus*. In order to achive this first goal, sequence data for the *tsf* gene and the neighbouring genes of the the bacteria phylogenetically related to *Geobacillus anatolicus* were used to design the appropriate oligonucleotide primers to amplify the chromosomal region carrying the *tsf* gene. This sequence information was then used to design primers to clone the complete *tsf* gene including 6 additional histidine residues at its C-terminal into an appropriate expression vector. This vector was used to express and purify the protein in high quantities in *Escherichia coli*. Ni-affinity chromotography was then used to purify the recombinant protein to homogeneity. *Geobacillus anatolicus* EF-Ts was found out to be fully active in binding *Escherichia coli* EF-Tu.

ÖZET

HİPERTERMOFİLİK *GEOBACILLUS ANATOLICUS* BAKTERİLERİNDEN REKOMBİNANT ELONGASYON FAKTÖRÜ TS'NİN KLONLANMASI, EKSPRESYONU VE SAFLAŞTIRILMASI

Geobacillus anatolicus elongasyon faktörü Ts His kuyruklu bir rekombinant protein olarak *Escherichia coli* içerisinde klonlanmış, ekpresyonu yapılmış ve saflaştırılmıştır. Bu çalışmanın ilk adımı *Geobacillus anatolicus tsf* geninin nükleik asit dizisinin belirlenmesidir. Bu amaç için, *Geobacillus anatolicus*'a filogenetik olarak en yakın bakterilerin *tsf* ve *tsf*'ye komşu genleri kullanılarak *tsf* genini taşıyan kromozomal bölgenin çoğaltılması amacıyla uygun oligonükleotit primerleri tasarlanmıştır. Daha sonra, bu dizi bilgisi kullanılarak, C terminal bölgesinde 6 ek Histidin amino asiti olan tam *tsf* genini uygun bir ekspresyon vektörüne klonlamak amacıyla uygun primerler tasarlanmıştır. Bu ekspresyon vektörü, proteinin *Escherichia coli* içinde yüksek miktarda ekspresyonunun sağlanması için kullanılmış ve proteinin saflaştırılması yapılmıştır. Rekombinant proteini homojen bir şekilde saflaştırabilmek amacıyla Ni-afinite kromatografisi kullanılmıştır. *Geobacillus anatolicus* EF-Ts proteininin *Escherichia coli* EF-Tu proteinine bağlanmada tam olarak aktif olduğu bulunmuştur.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS / ABBREVIATIONS

A	Adenine
A	Alanine
Å	Angstom
B	Asparagine
C	Cytosine
C	Cysteine
D	Aspartic acid
E	Glutamic acid
F	Phenylalanine
G	Guanine
G	Glycine
H	Histidine
h	Hour
I	Isoleucine
K	Lysine
L	Leucine
M	Methionine
N	Asparagine
P	Proline
Q	Histidine
R	Arginine
rpsB	ribosomal protein S2 gene

S	Serine
smbA	uridylate kinase gene
T	Thymine
T	Threonine
Tm	Melting temperature
tsf	elongation factor Ts gene
u	Unit
V	Valine
W	Tryptophan
Y	Tyrosine
Z	Glutamine

aa	Aminoacid
ADP	Adenosine 5’-diphosphate
AMP	Adenosine 5’-monophosphate
APS	Ammonium persulphate
Arg	Arginine
Asp	Aspartic acid
ATP	Adenosine 5’-triphosphate
bp	Base pair
BSA	Bovine serum albumin
Ca	Calcium
Cys	Cysteine
C-terminal	Carboxyl terminal
DNA	Deoxyribonucleic acid

dNTP	Deoxyribonucleotides
dsDNA	Double stranded deoxyribonucleic acid
DTE	1,4-Dithioerythritol
EDTA	Ethylenediaminetetraacetate
EF-G	Elongation factor G
EF-Ts	Elongation factor Ts
EF-Tu	Elongation factor Tu
EtBr	Ethidium bromide
GDP	Guanosine diphosphate
Gla	γ-carboxy glutamic acid
Gln	Glutamine
Glu	Glutamate
GTP	Guanosine triphosphate
His	Histidine
Ile	Isoleucine
IPTG	Iopropyl-1-thio-β-D-galactoside
kb	Kilo base
kDa	Kilo Dalton
LB	Luria-Bertani broth
Leu	Leucine
Lys	Lysine
Met	Methionine
Mg^{2+}	Magnesium
$MgCl_2$	Magnesium chloride
min	Minute

mg	Milligram
mM	Millimolar
mRNA	Messenger RNA
ng	Nano gram
Ni	Nickel
NTA	Nitrilotriacetic acid
N-terminal	Amino terminal
Nres	Nucleotide residues
PCR	Polymerase chain reaction
PEG	Polyethyleneglycol
pmol	picomole
PMSF	Phenylmethyl ulfonylfluoride
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
sec	Seconds
Ser	Serine
SDS	Sodiumdodecylsulphate
SDS-PAGE	SDS-polyacrylamide gel electropheresis
Taq	*Thermus aquaticus*
TE	Tris-EDTA
TEMED	N,N',N',N'-Tetramethylethylenediamine
Thr	Threonine
Tris	Tris(hydroxylmethyl)aminomethane
tRNA	Transfer RNA

Trp	Tryptophan
Tyr	Tyrosine
UV	Ultra Violet
Val	Valine

1. INTRODUCTION

Bacteria can be divided into several different classes based on the temperatures at which they grow best. Psychrophiles are bacteria whose optimum temperature is 15°C or lower but they can live in temperatures as low as -10°C. Mesophiles live accordingly at medium temperatures, 20-45°C, and as the temperature implies the human pathogens are included in this group. Thermophiles, in contrast, live above 45°C up to 100°C or more. *Geobacillus anatolicus* was isolated from a thermal vent in Balıkesir, at 98°C, and identified as a new hyperthermophilic bacteria in 2001 (Uysal *et al*., 2001).

Strains of thermophilic bacteria have been identified with optimum temperatures ranging from 55°C to 105°C. The strains that survive at extremely high temperatures are called extreme thermophiles, or hyperthermophiles, and have a temperature optimum of 80°C or higher. Thermophilic bacteria grow naturally in hot springs, tropical soils, compost heaps and in hot water heaters (both domestic and industrial) (Madigan and Martinko, 2005).

1.1. Protein Synthesis Machinery

Translational machinery, on which protein synthesis occurs, is composed of ribosomes, mRNAs, tRNAs and several protein factors. The ribosome is used as a platform for the assembly of mRNA, tRNAs and protein factors. Peptidyl-transferase activity resides in ribosomes (Marintchev and Wagner, 2005). Ribosome is composed of a large and a small subunit, referred to as 50S and 30S, respectively, in prokaryotes. The two subunits perform very different tasks during translation. The 30S subunit is responsible for the binding of mRNA and monitors the codon–anticodon interaction with tRNA. The large subunit interacts with the acceptor end of tRNAs to which the activated amino acids are attached. Peptidyl transferase center is in this subunit.. The two subunits therefore interact with opposite ends of the L-shaped tRNA and the entire ribosome displays three tRNA-binding sites. The A-site is where aminoacyl-tRNA (aa-tRNA) is accommodated on the ribosome when released by elongation factor Tu (EF-Tu, EF-1A in prokaryotes and eEF-

1A in eukaryotes) after GTP hydrolysis. Following peptidyl transfer, the elongated peptidyl-tRNA in the A-site is translocated to the P-site while deacylated tRNA in the P-site moves to the E-site. This movement of tRNAs is an intrinsic activity of the ribosome, yet it is greatly stimulated by elongation factor G (EF-G, EF-2 in prokaryotes and eEF-2 in eukaryotes) (Nilsson and Nissen, 2005). The sequence and structure of ribosomal RNAs, tRNAs, some of the ribosomal proteins and some of the translational factors are conserved in all kingdoms, which indicates the common origin of the translational apparatus. Translation can be divided into four distinct steps: initiation, elongation, termination and recycling (Marintchev and Wagner, 2005).

1.1.1. Translation Initiation

Translation initiation includes the steps between subunit dissociation after termination in the previous translation cycle and the assembly at an mRNA start codon of a ribosome for the next elongation cycle. During translation initiation, the initiator methionyl-tRNA, is assembled onto the P-site of the ribosome with mRNA, with the help of a set of initiation factors (IF's). Translation apparatus performs the following during initiation: (a) subunit dissociation and anti-association, (b) selection of the initiator tRNA, (c) selection of the correct translation start site, and (d) subunit joining at the start codon. At the end of initiation, the ribosome is ready to accept the first elongator tRNA and form the first peptide bond (Brock *et al.,* 1998).

1.1.2. Translation Elongation

Translation elongation phase is controlled by three elongation factors (EFs). Elongation factor EF-Tu (EF-Tu) forms a ternary complex with GTP and aminoacylated tRNAs (aa-tRNAs), protects the amino acid ester bond against hydrolysis and carries the aa-tRNA to the ribosomal A-site for the decoding of mRNA by codon-anticodon interactions. After ternary complex binding, the ribosome decodes the genetic information by a Watson–Crick-type interaction between the mRNA codon, which is exposed in the ribosomal A-site, and the anticodon of a cognate tRNA. Such a decoding event triggers GTP hydrolysis on EF-Tu, and EF-Tu bound to GDP is released from the ribosome (Andersen *et al.,* 2003). The nucleotide exchange factor elongation factor Ts (EF-Ts, EF-

1B in eukaryotes) catalyzes the recycling of EF-Tu·GDP into active EF-Tu·GTP (Clark *et al.*, 1999). EF-Tu consists of three domains. The first – domain I or the G domain – of 200

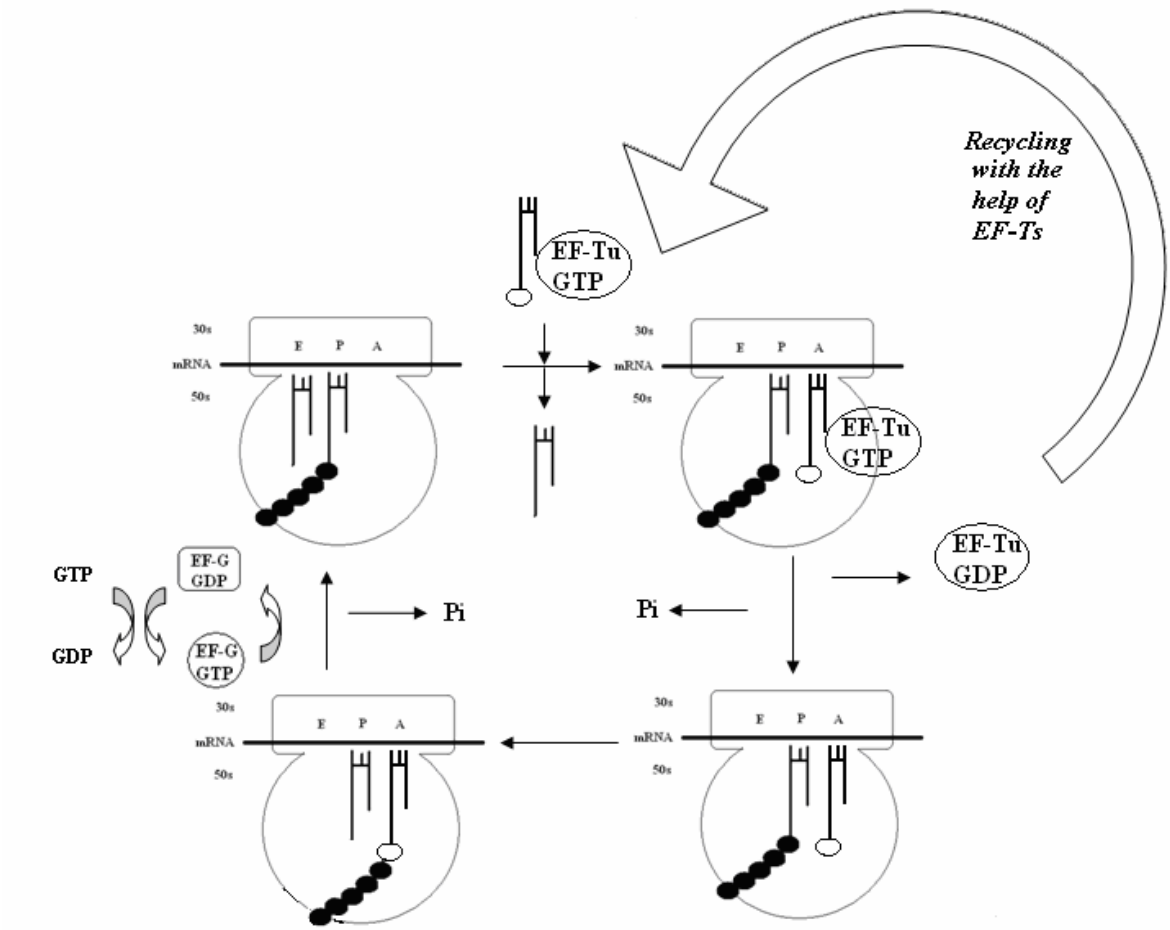

Figure 1.1. Protein synthesis elongation cycle in bacteria

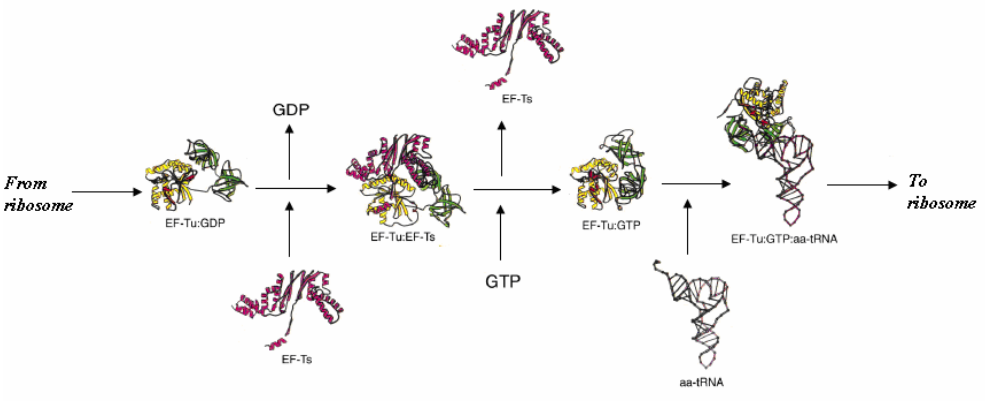

Figure 1.2. EF-Tu cycle (Partially adapted from Clark *et al.,* 1999)

residues is responsible for binding either GTP or GDP. This domain has the same basic structure as the G domains of all other G proteins. The two other domains – domain II and domain III – both contain about 100 residues and are both β barrels.

In all known structures, these two domains are held together in the same relative orientation and are, thus, likely to act as one rigid functional unit. EF-Tu undergoes a drastic conformational change between its active and inactive forms when bound to GTP and GDP, respectively. Also, switch 1 and switch 2 regions of the G domain of EF-Tu show significant local changes in their tertiary and secondary structures depending on the nucleotide bound. It is expected that this dramatic conformational change is a general characteristic of EF-Tu in all species. In support of this, an archaebacterial EF-Tu (aEF1A) – the sequence of which is more homologous to eEF1A than to EF1A – has been structurally determined in complex with GDP, and the result shows that its overall conformation is indeed similar to that of the bacterial EF-Tu–GDP structure (Vitagliano *et al.,* 2001). The large conformational change involved in re-activation of EF-Tu from the GDP form, as it comes off the ribosome and into the GTP form ready for interaction with a new aa-tRNA, needs EF-Ts as a catalyst (Andersen *et al.,* 2003).

The three-dimensional structure of the *Escherichia coli* EF Tu·EF-Ts complex was determined to a resolution of 2.5 Å (Kawashima *et al.,* 1996). From the crystal structure it appears that domain I of EF-Tu, the guanine nucleotide binding domain, interacts predominantly with the N-terminal half of EF-Ts, while EF-Tu domain III interacts with the C-terminal part of the EF-Ts core (i.e. residues 140±180 and 230±260). Mutational analysis of EF-Ts confirmed that residues in the N-terminal domain and in subdomain C of the EF-Ts core are crucial for this interaction (Zhang *et al.,* 1998). The driving force behind GDP release from EF-Tu is postulated to be the dislocation of Mg^{2+} from the molecule, which is catalyzed by EF-Ts (Hoogvliet *et al.,* 1999). After the release of EF-Tu·GDP, aa-tRNA is additionally selected by the proofreading step, and finally docked into the A-site where it is in contact with the peptidyl-tRNA in the ribosomal P-site. Peptide bond formation occurs leaving the deacylated tRNA in the P-site and peptidyl-tRNA in the A-site. At this stage, elongation factor EF-G (EF-2B in eukaryotes) in complex with GTP binds to the ribosome and catalyzes the translocation of tRNAs and mRNA on the ribosome (Clark *et al.,* 1999). After translocation, the mRNA moves by one codon carrying

the P-site bound deacylated-RNA to the E-site, and the A-site bound peptidyl-tRNA to the P-site.

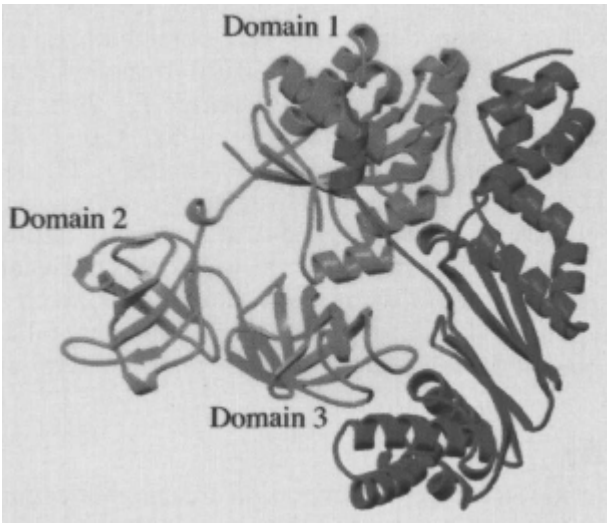

Figure 1.3. *Escherichia coli* EF-Tu·EF-Ts complex
(EF-Tu and its domains: light, EF-Ts: dark
(Adapted from Kawashima *et al.,* 1996))

1.1.3. Translation Termination and Recycling

Any one of the three codon triplets, UAA, UAG, or UGA, when present in the reading frame, codes for termination of protein synthesis. In eubacteria, two proteins, RF-1 and RF-2, are required to recognize the termination signal and to stimulate hydrolysis of the nascent peptidyl-tRNA by the peptidyl transferase of the 50S subunit. A third protein, RF-3, acts with GTP to eject these factors from the ribosome. *In vivo* and *in vitro* observations suggest that UGA can be read as termination, tryptophan, cysteine, or selenocysteine (Low and Berry, 1996). Also, UAA, UAG, or UGA can be suppressed without loss of viability. This misreading is necessary for the synthesis of some essential proteins (Tate and Brown, 1992). There is genetic and biochemical evidence that the code to terminate synthesis is longer than a single triplet and that, in certain cases, sequences in the mRNA either 5' or 3' of these codons may influence the decision of termination or insertion of an amino acid (suppression). Suppression is used for the synthesis of many bacteriophage and retroviral proteins. It has been proposed that the C-terminal end of the nascent peptide, which is specified by the two codons at the 5' side of UGA, markedly

influences this decision by promoting termination (Ganoza *et al.,* 2002). A ribosome-recycling factor, RRF, together with EF-G then completes the termination process by dissociating the two subunits. Eukaryotes do not have an RRF, but unlike the non-essential bacterial RF3, eRF3 is essential and could also be fulfilling the role of an RRF (Kisselev *et al.,* 2003).

1.2. Elongation Factor Ts

EF-Ts, the nucleotide exchange factor of EF-Tu, is present in the cell at about 0.2 molecule per ribosome which is 30 times less than its substrate; EF-Tu. EF-Ts associates with the EF-Tu·GDP complex and induces a structural change in EF-Tu that creates the necessary space for binding the γ-phosphate of the nucleotide, thus increasing the affinity for GTP with respect to that for GDP (Blumenthal *et al.,* 1977). It consists of two domains; EF-Ts domain and UBA/TS-N domain, which both interact with EF-Tu.GDP to catalyze the release of GDP from EF-Tu. EF-Ts can be divided into four structural units as follows: the N-terminal domain (residues 1–54), the core domain (residues 55–179 and 229–263), the dimerization domain (residues 180–228), and the C-terminal module (residues 264–282). The core domain is divided into subdomain N (residues 55–140) and subdomain C (residues 141–179 and 229–263). The N-terminal domain, subdomain N, and the C-terminal module all interact with domain I of EF-Tu, whereas subdomain C interacts with domain III (Zhang *et al.,* 1998).

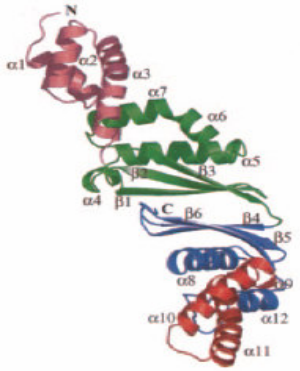

Figure 1.4. *Escherichia coli* EF-Ts
(aminoterminal domain: violet, subdomain N: green, subdomain C: blue and coiled coil motif: red (Adapted from Jeppesen *et al.,* 2005))

1.3. EF-Ts Gene Organization

The gene for EF-Ts, *tsf*, is expressed from a bicistronic operon, located at 4.1 min on the *Escherichia coli* map, together with ribosomal protein S2 (*rpsB*). A single major promoter immediately upstream of *rpsB* directs transcription of the *rpsB-tsf* operon. Two potential Rho-independent transcription terminators are present in the operon, one in the intergenic region between the two genes and one distal to *tsf*. Like EF-Tu and EF-G, EF-Ts synthesis is under stringent control and coordinated with ribosomal protein synthesis in a growth rate dependent manner (An *et al.,* 1981).

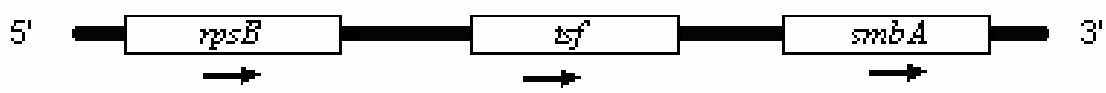

Figure 1.5. Gene organization of EF-Ts coding *tsf* gene

1.4. Protein Thermostability in Hyperthermophiles

As expected, the proteins which are produced by the thermophilic bacteria are more stable at high temperatures with respect to their non-thermophilic counterparts. The features which are commonly associated with elevated thermostability in proteins can be summarized as; relatively small solvent-exposed surface area, increased packing density that reduces cavities in the hydrophobic core, an increase in core hydrophobicity, decreased length of surface loops, and hydrogen bonds between polar residues. The enzymatic proteins lose their activities at high temperatures due to the loss of secondary and tertiary structures (Vetriani *et al*., 1998).

Protein stability is related to the function, and to the living temperature of the organism. Thermophilic and mesophilic proteins which are homologous to each other, can have similar stabilities at their respective organisms' optimum growth temperatures. When a mesophilic protein is exposed to higher temperatures, its stability decreases according to its stability curve. The sequence and structural variations keep the stability of the homologous thermophilic protein roughly unchanged, enabling it to function properly at the elevated temperatures. The control mechanisms of protein stability can be summarized

as; strengthening the hydrophobic core, reducing or eliminating surface loops and increasing and optimizing electrostatic interactions (Kumar and Nussinov, 2001).

Table 1.1 shows some of the higher temperature proteins. It can be concluded from the table that the dimerization domain of EF-Ts and EF-Tu/EF-Ts complex from *Thermus aquaticus* does no denature up to 95°C.

Table 1.1. Thermostability of various proteins
(Partially adapted from Kumar and Nussinov, 2001)

Protein family name	Organism	Stability	T_L (°C)	Nres
Citrate synthase	*Pyrococcus furtosus*	Half-life of 170 min at 100°C	100	741
Malate dehydrogenase	*Thermus flavus*	Fully active after 60 min at 90°C	70-75	644
Rubredoxin	*Pyrococcus furtosus*	Stable for >24h at 95°C	100	53
Cyclodextrin glucotransferase (GTAase)	*Thermoanaero-bacterium thermosufurigenes*	>90% catalytic activity when kept at 80°C for 5 h	60	683
EF-Tu and EF-Tu/Ts complex	***Thermus aquaticus***	**Temperature optimum ~70°C**	**70-72**	**405**
Dimerization domain of EF-Ts and EF-Tu/Ts complex	***Thermus thermophilus***	**Does not denature up to 95°C**	**70-75**	**284**
Glutamate dehydrogenase	*Pyrococcus furtosus*	Half-life of 12h at 100°C	70-105	2502
Lactate dehydrogenase	*Bacillus stearothermophilus*	Active for 30 min at 80°C	40-65	1264

2. PURPOSE

How the speed and the accuracy of translation can be maintained at high temperatures at which the Watson-Crick base pairing is extremely weakened, is not fully understood. Structural and functional information about thermophilic and hyperthermophilic translational apparatus is limited. In order to gain insight into translation process at high temperatures, translational factors that play important roles in protein synthesis must be well characterized. The aim of this thesis is to understand how the stability of the translational factors improves the overall stability of thermophilic protein synthesis. For this purpose, one essential protein synthesis elongation factor, EF-Ts will be studied. EF-Ts is the nucleotide exchange factor of EF-Tu, the elongation factor which carries aminoacylated tRNA's to ribosomes. *Geobacillus anatolicus*, a newly described hyperthermophilic bacteria isolated from a hot-spring at 98℃ in Turkey will be used as the source of EF-Ts. The *tsf* gene coding EF-Ts will be cloned, purified and studied functionally in its interactions with EF-Tu. This, in turn, can be used to set up an *in vitro* thermophilic protein synthesis machinery for excessive analysis of thermophilic protein synthesis.

3. MATERIALS

3.1. Bacterial Strains

3.1.1. *Geobacillus anatolicus* Strain

Geobacillus anatolicus was used as the source for Elongation factor Ts gene (*tsf*) in molecular cloning. *Geobacillus anatolicus* was previously collected from a hydrothermal vent at water temperature of 98ºC in Hisaralan, Balıkesir, Turkey and identified as a new species in our laboratory (GenBank Accession No: AF411064) (Uysal *et al*., 2001).

3.1.2. *Escherichia coli* Strain

Table 3.1. Genotype of the *Escherichia coli* strain used as the plasmid host throughout this study

JM1O9(DE3)	*end*A1, *rec*A1, *gyr*A96, *thi*, *hsd*R17 (rk^-,mk^+), *rel*A1, *sup*E44, *mcr*AΔ(*lac-pro*AB), [F′, *tra*D36, *pro*AB, *laq*[^{q}ZΔM15], λDE3

3.2. Oligonucleotide Primers

Table 3.2. Oligonucleotide primers used for the amplification of *tsf* gene and for DNA sequencing of EF-Ts
(D: A, G or T, R: A or G; Y: C or T; K: G or T; N: A, T ,G or C)

TsF45	5’- AAC WGG YGC DGG NAT GAT
TsR696	5’- RCG NCC TTC HAC CAT YTT
TsF592	5’- GAA GTG CCG CAA GAA GAG
TsR357	5’- CTT GGA CAG TGG AAC CGT
rpsBF	5’- ATY GTW GAY ACW AAC TGY GAY CC
smbAR	5’- CKC CAW AWR TTR CCD CCR C

Table 3.3. Oligonucleotide primers used for cloning of *tsf* gene

(M: A or C)

TsF2	5'- TTA AGA ACA TAT GGC GAT TAC MGC ACA AAT GGT AAA AGA G
TsR2	5'- AAT TAA ATT CGA ATT AAT GAT GAT GAT GAT GAT GTT GCT TTC TTA CTT GGC TCA TGA C

Table 3.4. Oligonucleotide primers used for plasmid DNA sequencing

T7 Forward	5'-TAA TAC GAC TCA CTA TAG GG
pRSET Reverse	5'-TAG TTA TTG CTC AGC GGT GG

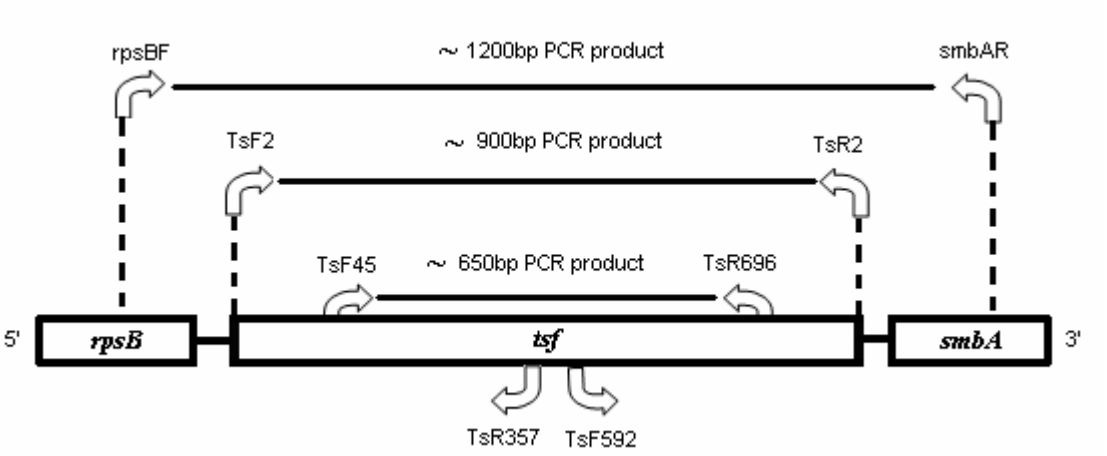

Figure 3.1. Primers used throughout this study shown together with their locations on schematical view of *tsf* gene and neighbouring two genes; *rpsB* and *smbA*

3.3. Enzymes Used Throughout This Study

Taq DNA polymerase (EC# 2.7.7.7) : Promega, USA

Nde I (EC# 3.1.21.4) : Promega, USA

Csp45 I (EC# 3.1.21.4) : Promega, USA

T4 DNA Ligase(EC # 6.5.1.1)	:	Promega, USA
Lysozyme (EC # 3.2.1.17)	:	Appligene, USA
RNase-A (EC # 3.1.27.5)	:	Promega, USA
Proteinase K (EC # 3.4.21.64)	:	Promega, USA

3.4. Chemicals

Agar	:	Oxoid, United Kingdom
Agarose	:	Sigma, USA
Ampicillin	:	Mustafa Nevzat, Turkey
Ammoniumpersulfate (APS)	:	Saveen, Sweden
Bromophenol Blue	:	Sigma, USA
Chloroform	:	Ambresco, USA
Coomassie Brilliant Blue R-250	:	Sigma, USA
Crystal Violet	:	Merck, Germany
D-Glucose	:	Riedel de-Häen, Germany
DNA molecular size markers	:	100 bp (Promega, USA) 1 kb (Promega, USA)
Dithioerythritol (DTE)	:	Fluka BioChemika, USA

Dimethyl sulfoxide (DMSO)	:	JT Baker, USA
Ethidium Bromide	:	Sigma, USA
Glacial Acetic Acid	:	Merck, Germany
Glass beads (0.1 mm diameter)	:	Biospec Products INC., USA
Glycerol	:	Merck, Germany
Guanosine 5'-diphosphate (GDP)	:	Promega, USA
HEPES	:	Sigma, USA
Imidazole	:	Sigma, USA
Isopropyl-ß-D-thiogalactopyranoside (IPTG)	:	Saveen, Sweden
β-Mercaptoethanol	:	Merck, Germany
Nitrilotriacetic acid (NTA)	:	Fluka, Switzerland
N,N,N',N'-Tetramethyl-ethylenediamine (TEMED)	:	Sigma, USA
Phenylmethylsulfonylfluoride (PMSF)	:	Sigma, USA
Phenol	:	Riedel de-Häen, Germany
Phenol Red	:	Sigma, USA
Polyethylene glycol (PEG 6000)	:	Merck, Germany

Protein Molecular Weight Markers	:	PageRuler Prestained Protein Ladder (Fermentas, USA) PageRuler Protein Ladder (Fermentas, USA)
Sodium Dodecyl Sulfate (SDS)	:	Kebolab, USA
Sucrose	:	Merck, Germany
Triton X-100	:	Merck,Germany
Tryptone	:	Oxoid, United Kingdom
Yeast Extract	:	Oxoid, United Kingdom

All other laboratory salts are analytical grade and obtained from Merck and Fluka, Germany.

3.5. Buffers and Solutions

3.5.1. Bacterial Growth

3.5.1.1. Stock Solutions for Growth Medium for *Geobacillus anatolicus*

5X Solution A	:	20 g/l yeast extract 40 g/l tryptone 15 g/l NaCl
10X Solution B	:	1 g/l NTA 0.6 g/l $CaSO_4.2H_2O$ 1 g/l $MgSO_4.7H_2O$ 1 g/l KNO_3 6.9 g/l $NaNO_3$

		1 g/l Na_2HPO_4
100X Solution C	:	28 mg/l $FeCl_3$
100X Solution D	:	220 mg/l $MnSO_4.H_2O$ 50 mg/l $ZnSO_4.7H_2O$ 1.6 mg/l $CuSO_4$ 2.5 mg/l $Na_2MoO_4.2H_2O$ 4.6 mg/l $CoCl_2.6H_2O$ 0.5 ml/l concentrated H_2SO_4

Final liquid culture medium (1X) contained appropriate dilutions of the concentrated stock solutions A, B, C, and D. After autoclaving D-glucose is added to a final concentration of 2%.

3.5.1.2. Growth Medium for *Escherichia coli* JM109 (DE3) Strain

1X LB-ampicillin Medium (pH 7.0)	:	5 g/l NaCI 10 g/l Tryptone 5 g/l Yeast Extract 150 µg/ml ampicillin
LB–Ampicillin Agar	:	5 g/l NaCI 10 g/l Tryptone 5 g/l Yeast Extract 14 g/l agar 150 µg/ml ampicillin

3.5.2. Stock Solutions for Genomic DNA Isolation

Homogenization Buffer	:	50 mM Tris-HCl (pH 8.0) 20 mM EDTA

20 % SDS	:	200 g/l SDS in H_2O
TE Buffer	:	10 mM Tris-HCl (pH 8.0) 1 mM EDTA

3.5.3. Stock Solutions for Polymerase Chain Reaction (PCR)

10 X PCR Buffer	:	10 mM Tris-HCl (pH 9) 50 mM KCl 0.1 % Triton-X 100 (Promega, USA)
$MgCl_2$	:	25 mM $MgCl_2$ (Promega, USA)
Deoxyribonucleotide mixture	:	100 mM of each dATP, dGTP, dCTP, dTTP (Promega, USA) in H_2O

3.5.4. Stock Solutions for Agarose Gel Electrophoresis

25 X TAE Buffer	:	121 g/l Tris (Base) 28.55 ml/l Glacial acetic acid 25 mM EDTA (pH 8.0)
1 or 2 % Agarose Gel	:	1 or 2 % (w/v) Agarose in 1 X TAE Buffer, containing 0.5 µg/ml Ethidium Bromide
6 X Loading Buffer	:	10mM Tris-HCl (pH 7.6) 0.03% bromophenol blue 0.03% xylene cyanol FF 60% glycerol 60mM EDTA

3.5.5. Stock Solutions for Enzyme Digestions

10X Buffer D for Nde I digestion	:	60 mM Tris-HCl (pH 7.9) 1.5 M NaCl 60 mM $MgCl_2$ 10 mM Dithiotheitol (DTT)
10X Buffer B for Csp45 I digestion	:	60 mM Tris-HCl (pH 7.5) 500 mM NaCl 60 mM $MgCl_2$ 100 mM DTT
10X T4 DNA Ligase Buffer	:	300 mM Tris-HCl (pH 7.8) 100 mM $MgCl_2$ 100 mM DTT 10 mM ATP

3.5.6. Stock Solutions for Competent Cell Preparation

TSS Buffer	:	10 % PEG 6000 5 % DMSO 20 mM MgSO4 5 g/l NaCl 10 g/l Tryptone 5 g/l Yeast extract

3.5.7. Stock Solutions for Transformation

SOC Medium	:	2 % Tryptone 0.5 % Yeast Extract 10 mM NaCl 2.5 mM KCl 10 mM $MgCl_2$

10 mM $MgSO_4$
20 mM D-Glucose

3.5.8. Stock Solutions for Plasmid DNA Isolation

TEL Solution	:	10 mM Tris-HCl (pH 8) 1 mM EDTA 5 mg/ml lysozyme
NAS Solution	:	0.2 M NaOH 1 % SDS
K-acetate Solution (pH 4.8)	:	60 ml/l 5M K-acetate 28.5 ml/l Glacial acetic acid
Resuspension Buffer	:	10 mM Tris-HCl (pH 8.0) 1mM EDTA 0.3 M NaCl

3.5.9. Stock Solutions for SDS-Polyacrylamide Gel Electrophoresis

Protein Lysis Solution	:	167 mM Tris-HCl (pH 6.8) 0.33 M SDS 10 % (w/v) sucrose 25 µl/ml β-mercaptoethanol 0.01 % (w/v) bromophenol blue
Solution I	:	20 % acrylamide 1 % N,N'-methylene-bisacrylamide (37.5/1)
Solution II A	:	3 M Tris (pH 8.9)

Solution II B	:	60 g/l Tris (base) 68.6 g/l $NaH_2PO_4.2H_2O$ (pH 7.8)
Gel Buffer (1:1)	:	200 ml Solution II A 8 ml 20 % SDS 1 ml TEMED 600 ml H_2O
Spacer	:	41.5 ml Solution I 31.25 ml Solution II B 1.25 ml 20 % SDS 0.25 ml TEMED 207 ml H_2O
5X SDS Sample Buffer	:	5 ml 20 % SDS 4 ml Solution II B 1 ml β-Mercaptoethanol 10 ml glycerol
10 % Ammoniumpersulfate	:	0.1 g/ml APS in H2O
40% Acrylamide Stock (37.5:1)	:	37.5 % Acrylamide 1 % N,N'-methylenebisacrylamide (purchased as a ready-made solution from Biorad, USA)
7X Solution 2C	:	24 g/l Tris (base) 115.2 g/l Glycine
Electrophoresis running buffer	:	167 ml/l 7X Solution 2C 5.83 ml/l 20 % SDS

Staining Solution	:	40 % Methanol 10 % Glacial Acetic acid 1 g/l Coomassie Brilliant Blue (R-250)
Destaining Solution	:	30 % Isopropanol 10 % Glacial acetic acid

3.5.10. Buffer for Cell Lysis and Homogenization

10X Homogenization buffer	:	500 mM Hepes-KOH (pH 7.5) 1.5 M KCl 100 mM $MgCl_2$ 10 % Glycerol 7 mM β-Mercaptoethanol 10 μM GDP 0.1 mM PMSF

3.5.11. Buffers for Ni-NTA Affinity Chromotography

Buffer A	:	50 mM Hepes KOH (pH 7.5) 0.15 M KCl 10 mM $MgCl_2$ 10 % Glycerol 7 mM β-Mercaptoethanol 10 μM GDP 0.1 mM PMSF
Buffer C	:	50 mM Hepes KOH (pH 7.5) 0.15 M KCl 10 mM $MgCl_2$ 10 % Glycerol

7 mM β-Mercaptoethanol
10 μM GDP
0.1 mM PMSF
1 M NH_4Cl
10 mM imidazole

3.5.12. Polymix Buffer for Dialysis, Storage and Assay

1X Polymix Buffer	:	95 mM KCl 5 mM NH_4Cl 5 mM $MgAc_2$. $4H_2O$ 0.5 mM $CaCl_2.2H_2O$ 8 mM Putrescine 1 mM Spermidine 5 mM K-phosphate pH 7.5 1 mM DTE

10 X Polymix buffer is prepared without K-phosphate to avoid precipitation and DTE. In assays, correct working strength of 1X Polymix is obtained by mixing 10X Polymix buffer appropriately with 100 mM K-phosphate and 50 mM DTE. For dialysis DTE is freshly added as solid.

3.6. Expression Vector

pCR® T7/NT-TOPO®	:	INVITROGEN, USA
pT7D3	:	Derived from pCR®T7/ NT-TOPO® after Nde I and Csp45 I restriction endonuclease reactions (this work)
pT7D3Ts	:	Derived from pT7D3

after ligation of PCR fragment of 6XHis-tag containing *Geobacillus anatolicus tsf* gene (this work)

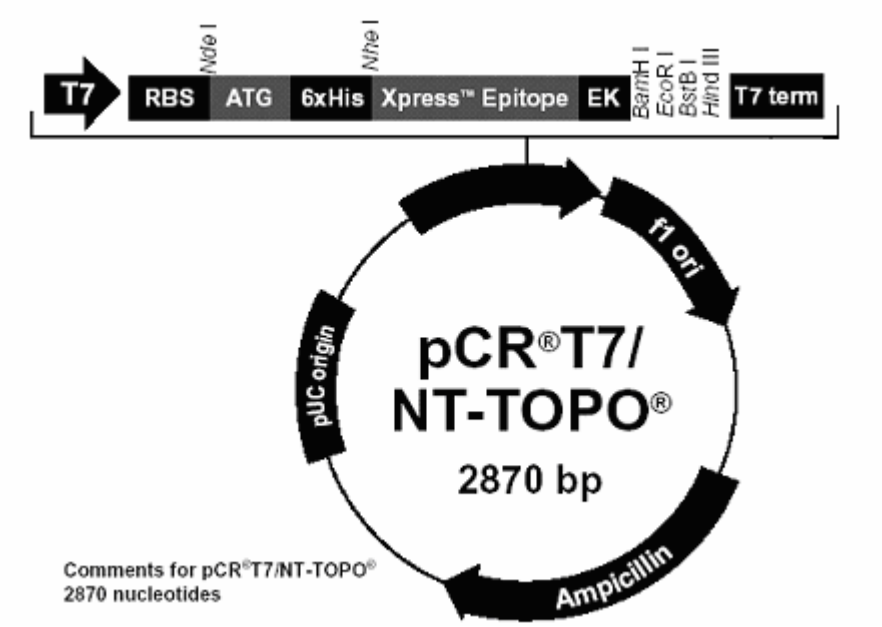

Figure 3.2. Map of pCR® T7/NT-TOPO® vector
(From http://www.invitrogen.com/content/sfs/vectors/pcrt7nttopo_map.pdf)

3.7. PCR Product Purification

Wizard SV Gel and PCR Clean-Up System : Promega, USA

3.8. Equipments

Autoclave : Model MAC-601 (Eyela, Japan)

Balance : MonoBloc Model PB602, USA

Beat Beater : Model 1107900, Biospec

		Products, USA
Centrifuges	:	Beckman Microfuge E,Germany Beckman L7-55 Ultracentrifuge, USA Beckman J2-MC Centrifuge, USA Beckman J2-21 Centrifuge, USA Centurion K2R, United Kingdom
Deep Freezers (-20°C)	:	Arçelik 2031D and 2020D, Turkey
Deep Freezer (-80°C)	:	Sanyo Ultra Low, UK
Documentation System	:	BioRAD, USA
Gel Drier	:	Honeywell, UK
Heat-block	:	Grant FE10, Cambridge, UK
Incubator	:	65°C (LEEK, USA)
Liquid Chromotography system	:	Pharmacia, Sweden
Magnetic Stirrer	:	Jankey δ Kunkel KM02, USA
Ovens	:	Microwave Oven (Vestel, Turkey) 110°C (Gallenkamp 300, UK)
pH Meter	:	Jenway 3010, USA
Power Supplies	:	Betherde Research Model 200 (BMC, Uppsala)
Refrigerator(4°C)	:	Arçelik 1071D, Turkey

Rotors	:	Beckman JS-7.5A, JA-14, Germany Beckman 70 Ti, USA
Spectrophotometer	:	UV-Visible Agilent 8453, USA
Thermocyclers	:	Techne (Progene, UK) Techne Gradient (BioRAD, USA)
UV Transilluminator	:	Chromato-Vue Transilluminator, Model 1TM-20UVP (USA)
Water Baths	:	RE100B, Grant, UK and Innova 3100, USA
Vortex	:	MS2 Minishaker IKA, USA

4. METHODS

4.1. Extraction of the Genomic DNA from *Geobacillus anatolicus*

Geobacillus anatolicus cells taken from -80°C glycerol stock were streaked on agar plate containing the growth medium at pH 8.5 and incubated at 65°C overnight. A toothpick of cells from this plate was re-suspended in 480 µl TE buffer containing 15 µl of 20 % SDS and 6 µl of 10 mg/ml proteinase K in a sterile eppendorf tube. After vortexing, the mixture was incubated at 37°C for 1 h. One volume of TE-saturated phenol: chloroform mixture (1:1) was added to the sample to deproteinize DNA. The mixture was then vortexed and centrifuged for 3 min at room temperature in microcentrifuge to obtain a clear solution. The upper phase containing DNA was transferred into a sterile eppendorf tube. Extraction steps were repeated three times. The upper phase from the last phenol extraction step was mixed with one volume of chloroform. After vortexing and centrifugation, the upper phase was transferred into a new sterile eppendorf tube and 0.1 volume of 3M sodium acetate (pH 5.2) and then 0.6 volume of isopropanol were added. Samples were left for 10 min at room temperature prior to centrifugation for 3 min at 4°C. Precipitated DNA was washed with two volumes of 70 % ice-cold ethanol in order to remove salts. After air-drying the precipitated DNA, 700 µl TE buffer was added to dissolve DNA. RNase-A was then added to a final concentration of 0.1 mg/ml and the sample was incubated at 37°C for 30 min in order to remove RNA. To remove RNase-A, one volume of water saturated phenol was added to the sample. After phenol extraction, and centrifugation, the upper phase was transferred into a sterile eppendorf tube. Extraction step was repeated twice with chloroform. The upper phase over the chloroform extraction was transferred to a clean sterile tube and 0.1 volume of 3M sodium acetate (pH 5.2) and 0.6 volume of isopropanol were added and left for 10 min at room temperature to precipitate DNA. DNA was collected by centrifugation for 3 min at 4°C. The supernatant was discarded and the DNA pellet was washed with two volumes of 70 % ice-cold ethanol, air dried and dissolved in 100 µl TE buffer. DNA was left overnight at 4°C and then stored at -20°C until use. The concentration of the genomic DNA was determined spectrophotometrically. The calculation was based on the fact that 50 µg of double

stranded DNA has an absorbance of 1.0 at $OD_{260\ nm}$. 260 nm/280 nm ratio of the absorbances are used for DNA quality estimations.

4.2. Analysis by Agorose Gel Electropheresis

Genomic DNA was electrophoretically analyzed on 1 % agarose gel in TAE buffer. TAE buffer (40 ml) containing 1 % agarose was melted in microwave, cooled down to 55°C and ethidium bromide was added to a final concentration of 0.5 µg/ml under a fume hood. The solution was poured onto the eletrophoresis plate and left to polymerize at room temperature, then placed into an electrophoresis chamber containing TAE buffer. Then the comb was removed. After mixing 5 µl of samples with 1 µl of 6X loading dye, the samples were applied to wells. A molecular size marker was also included in the run. The electrophoresis was at 150 Volt for 15-20 min. The gels were then analyzed under UV_{254nm} light and the images were recorded digitally.

4.3. Analysis of *Geobacillus anatolicus* EF-Ts Gene

4.3.1. PCR Amplifications of *Geobacillus anatolicus* EF-Ts Gene

Geobacillus anatolicus tsf gene was amplified with primer pairs containing both degenerate and non-degenerate bases. The primer pairs used in PCR reactions performed for obtaining sequence data are TsF45/ TsR696, rpsBF/smbAR and they are used to amplify partial and full sequences from the EF-Ts gene, respectively. TsF2/TsR2 PCR reaction was used for cloning of the *tsf* gene. PCR reactions were performed in 25 µl volume in PCR buffer containing 1.5 mM of $MgCl_2$, 0.2 mM of each dNTP, 1 pmol of each primer, 50 ng of genomic DNA and 0.25 units of *Taq* DNA polymerase. PCR cycles were as follows:

94°C for 2 minutes 30 seconds
95°C for 1 minutes 30 seconds (denaturation)
50-60°C gradient for 1 minute 30 seconds (annealing) } X 35
72°C for 2 minutes (extension)
72°C for 7 minutes

After amplification, 5 µl of each PCR product was mixed with 1 µl 6 X loading dye and electrophoretically analized on 1 % agarose gel.

4.3.2. Purification of the PCR Products

PCR products were purified by using Wizard SV PCR Clean-Up system (Promega) following the instructions given by the supplier. Namely, an equal volume of membrane binding solution was added to the PCR reaction mix, vortexed briefly and applied onto a SV minicolumn assembly (a minicolumn which was placed on a 2 ml collection tube) and incubated at room temperature for 1 min. After centrifugation for 1 min at 5000xg, the flowthrough was discarded and the minicolumn was reinserted into a new 2 ml collection tube. The column was washed with 700 µl membrane washing solution. After centrifugation for 1 min at 5000xg, the flowthrough was discarded and the column was inserted into a new 2 ml collection tube. The washing step was repeated by adding 500 µl membrane washing solution and centrifugation for 5 min at 5000xg. The flowthrough was discarded and a new sterile 2 ml collection tube was placed under the minicolumn. After adding 50 µl nuclease-free water to the minicolumn, the column was left at room temperature for 1 min then centrifuged for 5 min at 5000xg to collect the purifed DNA. Purified PCR products were electrophoretically analized on 1 % agarose gel.

4.3.3. DNA Sequencing

In addition to their uses in PCR reactions, the primer pair TsF45/TsR696 was used for the sequencing to get the primary information about the *Geobacillus anatolicus tsf* gene. Also TsF592/TsR357 and rpsBF/smbA primer pairs were used for determination of full *tsf* gene sequence. DNA sequencing was carried out by cycle sequencing method using DYEnamic ET Terminator Cycle Sequencing Kit (Amersham Biosciences, USA) using ABI 9700 Thermocycler (Applied Biosystems) at Acıbadem Hospital, Istanbul. Purified PCR products (or plasmids) were mixed with the sequencing mixture containing 5 pmol primer, 8 µl of sequencing reagent premix and water to a total reaction volume of 20 µl. The sequencing reaction was performed for both forward and reverse primers separately. The reaction conditions are 25 cycles of 95^{o}C for 20 s, 50^{o}C for 15 s and 60^{o}C for 60 s. After cycle sequencing, the unbound dyes were removed by using DyeEx 2.0 Dye removal

Kit (Qiagen). The purified products were analysed on the ABI 3100 Genetic Analyzer (Applied Biosystems). The sequencing results were analyzed by using CHROMAS software.

4.4. Cloning of *Geobacillus anatolicus tsf* Gene

4.4.1. Restriction Enzyme Digestion of PCR Fragment and Plasmid pCR® T7/NT-TOPO®

The restriction enzyme digestions were done in a total of 40 µl reaction volume. First reaction tube of 20 µl contained Buffer D and acetylated BSA with 2µg of either PCR fragment (resulting from TsF2/TsR2) or plasmid and 20u of Nde I restriction enzyme. These mixtures were incubated at 37ºC for 4h. Second reaction was done by direct addition of Buffer B and acetylated BSA and 40u of Csp45 I and completed to 40µl with dH_2O in both reaction tubes. These new reaction mixes were incubated at 37ºC for 4h. As a control for the digestions, 2µg of plasmid pCR® T7/NT-TOPO® was digested in a reaction tube of 20 µl contained Buffer B and acetylated BSA with 10u of Csp45 I at 37ºC for 2h. 5µl of the digested samples in parallel with the undigested plasmid were analyzed electrophoretically on 1% agarose gel for confirmation of digestion. pCR® T7/NT-TOPO® was named as pT7D3 after successful double-digestion.

4.4.2. Purification of the Digested PCR Fragment and Plasmid pT7D3

The PCR fragment digestion could not be confirmed until the ligation of the fragment into the cloning vector and transformation of the cells with this vector, because the digested part is very short in base pairs to be observed on an agarose gel. To get rid of the digested small fragments resulting from both enzyme digestions, the digestion reaction of PCR fragment was fully loaded on a %1 agarose gel. Small fragments (~9bp) migrate much faster than the longer fragments (~900bp), the longer band was cut out of the gel and this gel slice was purified using Wizard SV Gel and PCR Clean-Up System, Promega, USA. The 40µl plasmid digestion reaction mix was fully loaded on a %1 agarose gel. After the run, the longer of the two bands appearing in the gel was cut out of the gel. This gel slice was applied to Wizard SV Gel and PCR Clean-Up System, Promega, USA for the

extraction of the digested DNA from gel. Restriction enzyme positive control reaction of 15µl with known digestibility of restriction enzymes was also purified using Wizard SV Gel and PCR Clean-Up System, Promega, USA.

4.4.3. Ligation of the Digested PCR fragment and Plasmid pT7D3

The ligations were carried out in a total volume of 10µl in T4 DNA Ligase Buffer, 100ng of digested plasmid (pT7D3), 1 u of T4 DNA Ligase and varying amounts of digested PCR fragments from 10ng to 400ng. The ligations were carried out at 11ºC, overnight.

As a control for ligation reaction the Csp45 I digested and purified plasmid was used. Control ligation reaction was carried out in 30µl containing T4 Ligase Buffer, 700ng digested plasmid and 7u of T4 DNA ligase. This reaction was incubated at 11ºC, overnight. After successful ligation the new plasmid was named as pT7D3Ts.

4.5. Transformation of the *Escherichia coli* JM109 (DE3) Cells

4.5.1. Preparation of *Escherichia coli* JM109 (DE3) Competent Cells

Escherichia coli JM109(DE3) cells (Table 3.1) were grown overnight at 37^{o}C in liquid LB medium with orbital shaking at 200 rpm. A fresh culture was prepared by taking 200 µl overnight culture into 20 ml liquid LB medium, and growth continued at 37^{o}C. The cell growth was monitored spectrophotometrically by taking 1 ml samples and measuring the OD_{600} values. When the absorbance reached to OD_{600} = 0.5, the cells were transferred onto ice to cool and then harvested by centrifugation at 4^{o}C for 10 min at 5000 rpm using Beckman JA-14 rotor. After discarding the supernatant, 2 ml ice cold TSS buffer was added to the cell pellet, and the sample was kept on ice until cells were gently resuspended. This suspension of competent cells were aliquoted into pre-cooled eppendorf tubes on ice in the cold room. The samples (each 100 µl) were frozen in liquid nitrogen for 2 min, then stored at –80^{o}C until use.

4.5.2. Transformation

Competent *Escherichia coli* JM109(DE3) cells (100 µl) were thawed on ice and 4 µl of ligation mix dilutions were added and the samples were incubated on ice for 30 min. As positive controls, 4 µl of the Csp45 I digested, purified and ligated plasmid mix and also 4 µl of non-processed plasmid were used in transformations. As a negative control 4 µl H_2O was used instead of the plasmids. The cells were heat shocked at 42^oC for 1 min 30 sec, then immediately transferred onto ice and kept on ice for 2 min. Pre-heated 1 ml SOC medium was then added to each sample. The samples were incubated at 37^oC for 1 h 30 min. For each transformation 1 ml of each sample was layered on LB-agar plates containing 150 µg/ml ampicilin, and incubated overnight at 37^oC. The number of colonies appeared the next day on each plate was recorded. Individual colonies were re-streaked at least twice in order to obtain homogeneous populations.

4.6. Confirmation of the Plasmid pT7D3Ts, Carrying the *tsf* Gene

4.6.1. Plasmid DNA Purification

A broad sweep of cells was taken from the transformants which were grown on an agar plate after an overnight incubation. Cells were resuspended in 200 µl TEL solution by vortexing vigorously. 400 µl NAS solution was then added and mixed by inverting the tube several times. This mixture was kept on ice for 5 min. 300 µl of KAC solution was added and mixed by inverting and the mixture was kept on ice for 15 min. The mixture was centrifuged for 4 min at 20000xg. Approximately 750µl of the supernatant was transferred to a clean eppendorf tube and equal volume of ice cold 100% isopropanol was added and the mixture was kept at -70°C for 15 mins for the precipitation of plasmid DNA. Then, the mixture was centrifuged for 10 min at 20000xg at +4°C and the supernatant was discarded. The DNA pellet was washed with 1ml ice cold 70% ethanol and centrifuged again for 2 min at 20000xg. The supernatant was aspirated and the DNA pellet was air-dried for 15 min, dissolved in 50 µl TE buffer and stored at –20^oC.

Purified plasmid DNA samples were analyzed electrophoretically on 1 % agarose gel in TAE buffer (150 V, 30 min) alongside with a 1 kb DNA ladder as a size marker.

4.6.2. Nde I and Csp45 I Double-digestions of the Purified Plasmid pT7D3Ts

The restriction enzyme digestions were done in a total of 40 µl reaction volume. First digestion tube contained (in 20 µl) 1X Buffer D and 10ng/µl acetylated BSA with 2µg of purified plasmid and 20u of Nde I restriction enzyme. Nde I digestion was at 37ºC for 4h. Second reaction was done by direct addition of 1X Buffer B and 10ng/µl acetylated BSA and 40u of Csp45 I to the first digestion mix and the volume was completed to 40µl with H_2O. Csp45 I digestion was also at 37ºC for 4h. Reaction was stopped by heating the reaction mixture at 80ºC for 15 mins.

After digestion, samples were anayzed electrophoretically on 1 % agarose gel in TAE buffer.

4.6.3. Control PCR of the Purified Plasmid pT7D3Ts

PCR was carried out with the primers TsF2/TsR2 in order to confirm that the fragment carrying the *tsf* gene is inserted in the purified plasmid. PCR reactions were performed in 25 µl volume in PCR buffer containing 1.5 mM of $MgCl_2$, 0.2 mM of each dNTP, 1 pmol of each primer, 50 ng of purified plasmid DNA and 0.25 units of *Taq* DNA polymerase. PCR cycles were as follows:

94^{o}C for 2 minutes 30 seconds
95^{o}C for 1 minutes 30 seconds (denaturation)
57^{o}C for 1 minute 30 seconds (annealing) } X 30
72^{o}C for 2 minutes (extension)
72^{o}C for 7 minutes

After amplification, 5 µl of each PCR product was mixed with 1 µl 6 X loading dye and electropheretically analized on 1 % agarose gel.

4.6.4. PCR Amplification of the Plasmid pT7D3Ts for DNA Sequencing

Instead of direct sequencing of the plasmids to confirm whether they contained the *tsf* gene, this region of the plasmid is amplified by PCR by using the primer pair T7 Forward/ PR-Set Reverse. PCR reaction was performed in 100 µl volume in PCR buffer containing 1.5 mM of $MgCl_2$, 0.2 mM of each dNTP, 1 pmol of each primer, 200 ng of purified plasmid DNA and 1.25 units of *Taq* DNA polymerase. PCR cycles were as follows:

94°C for 2 minutes 30 seconds
95°C for 1 minutes 30 seconds (denaturation)
52°C for 1 minute 30 seconds (annealing) } X 30
72°C for 2 minutes (extension)
72°C for 7 minutes

After amplification, 5 µl of PCR product was mixed with 1 µl 6 X loading dye and electropheretically analized on 1 % agarose gel.

4.6.5. Long Term Storage of Transformed Bacteria

Transformants of *Escherichia coli* JM109(DE3) cells harboring the *Geobacillus anatolicus tsf* gene were grown overnight at 37°C on an LB-plate containing 150 µg/ml ampicilin. A broad sweep of cells taken by a toothpick was resuspended in LB medium containing 50 % glycerol. Cells were resuspended until a homogeneous suspension was obtained and stored at -80°C.

4.7. Expression of the *tsf* Gene

Escherichia coli strain JM109(DE3) carrying the plasmids with (pT7D3Ts) and without (pCR® T7/NT-TOPO®) the inserted *tsf* gene were grown in 10 ml of LB medium containing 150 µg/ml ampicilin at 37°C with orbital shaking at 200 rpm until the OD_{600} reaches approximately the absorbance of 0.6. Then, 500 µl aliquots were taken to

eppendorf tubes, and placed on ice. These samples were later used as controls to compare the expression levels between the uninduced and induced cells.

Isopropylthiogalactopyranoside (IPTG) was added to the remaining cultures to a final concentration of 2 mM. Aliquots of 500 µl were taken after 1 h, 2 h and 4 h into eppendorf tubes. Cells were collected by centrifugation in microcentrifuge at 4^{o}C, then supernatants were aspirated. The cell pellets were resuspended in 100 µl protein lysis solution and stored at –20^{o}C until SDS-polyacrylamide gel electropheresis (SDS-PAGE) analysis.

Before SDS-PAGE analysis, samples were incubated for 2 min at 85^{o}C in a water-bath for heat denaturation of the proteins, then 15 µl of each sample or 10 µl protein molecular weight markers were applied to 10 % SDS-PAGE gel. Electrophoresis was at 30 mA for 3 h. After electrophoresis, the gel was stained with Coomassie Brilliant Blue staining solution for minimum 2 h then destained until the protein bands became visible.

4.8. Purification of *Geobacillus anatolicus* Elongation Factor Ts

Escherichia coli JM109(DE3) cells harboring the cloned expression vector pT7D3Ts were grown overnight at 37^{o}C in LB plate containing 150 µg/ml ampicilin. A fresh culture was prepared by inoculating 1600 ml of LB (4X400 ml, in different culture tubes) containing 150 µg/ml ampicilin. IPTG was added to a final concentration of 2 mM when OD_{600} of the culture reached the absorbance of 0.6, and growth was continued for another 4 h. The cells were harvested by centrifugation at 10000 rpm at 4^{o}C for 15 min in Beckman JS-14 rotor. The cell pellet was kept at –80^{o}C until use.

For recombinant protein purification, total 12 g of cell pellet was resuspended in 12 ml of ice-cold homogenization buffer. After adding 24 ml of glass-beads (0.1 mm in diameter), the cell mixture was transferred to the mixing chamber of the bead-beater. Ice was placed around the chamber to prevent warming. Cells were disrupted 3 times each for 30 sec, in the bead-beater. Between each run, 5 min intervals were given in order to prevent heating. The cell lysate was centrifuged at 16000 rpm at 4^{o}C for 20 min using

Beckman JA-20 fixed-angle rotor. The supernatant was tranferred to a clean centrifuge tube and centrifuged at 45000 rpm at 4^{o}C for 2 h. The supernatant was saved and then applied directly onto Ni^{+2}-affinity column.

4.9. Ni-Affinity Column Chromatography

Ni-NTA chromatography material (Pharmacia) was pre-equilibrated with 50 ml of equilibration buffer containing 10 mM imidazole and packed into a glass column (column dimensions: 0.9x6 cm). The supernatant of the cell lysate was applied to the column with a speed of 2 ml/min. Unbound proteins were washed for about 50 ml of equilibration buffer until OD_{260nm} reached the base-line. His-tagged recombinant *Geobacillus anatolicus* Elongation Factor Ts was eluted with a 100 ml linear gradient from 10 to 250 mM imidazole in elution buffer. Fractions of 5 ml were collected.

Aliquots of 20 µl from fractions were mixed with 5 µl 5X SDS-PAGE sample buffer, denaturated for 2 min at 85^{o}C then applied to SDS-PAGE along with a molecular weight marker. Electrophoresis was for 4 h at 30 mA.

4.10. Protein Dialysis

The fractions from Ni-affinity column containing the recombinant EF-Ts, as judged by SDS-PAGE, were pooled and concentrated by ammonium sulphate precipitation (0.5 g/ml). Precipitated protein was dissolved in a minimum volume of polymix buffer and dialyzed overnight at 4^{o}C against 1lt polymix buffer. Dialysis buffer was changed after 4 h and dialysis continued for another 10 h. After dialysis, the protein sample was centrifuged for 10 min in a microfuge to clarify the sample from particles. EF-Ts was divided into 100 µl aliquots and stored at -80^{o}C.

4.11. Determination of Protein Concentration

EF-Ts concentration was determined by Bradford assay (Bradford, 1976). A series of tubes containing Bovine Serum Albumin (BSA) from 1 µg to 10 µg were prepared in a final volume of 100 µl by appropriate dilutions of 1 mg/ml stock solution of BSA in 0.15

M NaCl. A series of tubes containing the protein sample were also prepared in a final volume of 100 µl by diluting the protein sample in 0.15 M NaCl. After adding 1 ml Bradford's reagent to each tube, samples were vortexed and left at room temperature for 10 min., then the absorbance of each sample at OD_{595nm} was measured. A plot of the absorbance at 595_{nm} versus the volume of the protein sample was obtained. Similarly a plot of the absorbance at 595_{nm} versus the amount (µg/ml) of the BSA protein was obtained. The unknown protein concentration was calculated using BSA as a standard.

4.12. Binding of *Geobacillus anatolicus* EF-Ts to *Escherichia coli* EF-Tu

Binding of recombinant *Geobacillus anatolicus* EF-Ts to *Escherichia coli* EF-Tu was analyzed by their positional gel shifting on 10% non-denaturing PAGE. Titrations of recombinant *Geobacillus anatolicus* EF-Ts in the presence and absence of 300 pmol *Escherichia coli* EF-Tu were prepared in 50 µl in 1X Polymix buffer. In a parallel experiment *Escherichia coli* EF-Ts were used with *Escherichia coli* EF-Tu. Samples were prepared on ice and then incubated at 37ºC for 15 min. After incubations samples were kept on ice until use. 20 µl samples were mixed with 5 µl loading dye (phenol red in 50% glycerol) from which 20 µl were applied to 10% non-denaturing PAGE gel in 8mM Tricine pH 8.2. Electrophoresis was done in a cold room at 4ºC in the same buffer. Before electrophoresis of the samples a prerun for 1 hr at 250 V was performed. Samples were then applied and electrophorized for 3 hrs at 250V. Proteins were visualized after Comassie staining.

5. RESULTS

5.1. Identification of the *Geobacillus anatolicus* Elongation Factor Ts Sequence

5.1.1. Constructing Degenerate Primers for *tsf* Gene

At the start of this study, there was no genomic sequence or corresponding amino acid sequence data available for the *Geobacillus anatolicus tsf* gene. In order to obtain this information the genomic sequences of the *tsf* genes from various bacteria which are closely related to *Geobacillus anatolicus* were used. The conserved regions in the *tsf* genes were used for designing degenerate primers. (See Table 5.1, Figure 5.1 and Figure 3.1)

Table 5.1. Bacterial *tsf* genes used for primer designation

Source organism	**Gene ID**	**Nucleic acid sequence length**	**Aminoacid sequence length**	**Molecular Weight (kDa)**
Bacillus licheniformis ATCC 14580	3031063	882 bp	293	32
Bacillus halodurans C-125	891042	882 bp	293	32,3
Bacillus subtilis subsp. subtilis str. 168	939601	882 bp	293	32,2
Bacillus cereus E33L	3026190	888 bp	295	32,3
Bacillus anthracis str. Sterne	2848103	888 bp	295	32,3
Staphylococcus aureus subsp. aureus MSSA476	2861932	882 bp	293	32,4
Listeria monocytogenes str.4b F2365	2798556	885bp	294	32,6
Oceanobacillus iheyensis HTE831	1017975	885bp	294	32,6

```
          ....|....| ....|....| ....|....| ....|....| ....|....|
              5         15         25         35         45
B lic     ATGGC---CA ATCCTGCACA AATGGTAAAA GAGCTTCGCG AAAA**AACTGG**
B hal     ATGGC---AA TTACGGCAAG CATGGTAAAA GAGCTGCGCG AAAA**AACAGG**
B sub     ATGGC---AA TTACTGCACA GCAAGTAAAA GAACTGCGTG AAAA**AACTGG**
B cer     ATGGC---AA TCACTGCACA AATGGTAAAA GAATTACGTG AAAA**AACTGG**
B ant     ATGGC---AA TCACTGCACA AATGGTAAAA GAATTACGTG AAAA**AACTGG**
S aur     ATGGCAACTA TTTCAGCAAA ACTTGTTAAA GAATTACGTG AAAA**AACTGG**
L mon     ATGGCTAATA TTACAGCTCA AATGGTAAAA GAATTACGCG AAAA**AACTGG**
O ihe     ATGGC---TA TTACTGCACA AATGGTTAAA GAATTACGTG AAAA**AACTGG**
Conse     ATGGCWAMHA WYHCDGCWMR VMWDGTWAAA GARYTDCGYG AAAA**AACWGG**

          ....|....| ....|....| ....|....| ....|....| ....|....|
             55         65         75         85         95
B lic     **CGCAGGCATG ATG**GACTGCA AGAAAGCGTT AACGGAAACT GACGGCAACA
B hal     **CGCAGGGATG ATG**GATTGTA AAAAAGCATT AACAGAAACG AACGGAGATA
B sub     **CGCGGGCATG ATG**GATTGTA AAAAAGCGTT AACTGAAACT GACGGAGATA
B cer     **CGCAGGTATG ATG**GACTGCA AAAAAGCTTT AACAGAAACT AACGGCGACA
B ant     **CGCAGGTATG ATG**GACTGCA AAAAAGCTTT AACAGAAACT AACGGCGACA
S aur     **CGCGGGTATG ATG**GATTGTA AAAAAGCGCT AACTGAAACT GATGGTGACA
L mon     **TGCTGGTATG ATG**GATTGTA AAAAAGCACT TGTAGAAACA GAAGGAGATA
O ihe     **TGCTGGAATG ATG**GATTGTA AAAAAGCACT ACAAGAAACG AATGGTGATA
Conse     **YGCDGGNATG ATG**GAYTGYA ARAAAGCDYT WVHDGAAACD RAHGGHRAYA

          ....|....| ....|....| ....|....| ....|....| ....|....|
            105        115        125        135        145
B lic     TGGATAAAGC GATCGATCTT CTAAGAGAAA AAGGAATCGC AAAGGCTGCC
B hal     TGGATAAAGC GATTGACTAT CTTCGTGAAA AAGGGATTGC GAAAGCGGCG
B sub     TGGACAAAGC AATTGACCTT TTAAGAGAAA AAGGAATTGC AAAAGCAGCG
B cer     TGGAGAAGGC AATTGACTTC TTACGTGAAA AAGGTATCGC GAAAGCTGCT
B ant     TGGAGAAGGC AATTGACTTC TTACGTGAAA AAGGTATCGC GAAAGCTGCT
S aur     TCGATAAAGC GATTGATTAC CTACGTGAAA AAGGTATTGC TAAAGCAGCT
L mon     TGGAAAAAGC AATTGACTAT CTTCGTGAAA AAGGAATCGC TAAAGCTGCG
O ihe     TTGAACAGGC AATTGATTTT CTTCGTGAAA AAGGAATGGC GAAAGCTGCG
Conse     TBGANMARGC RATYGAYYWY YTWMGWGAAA AAGGDATBGC DAARGCDGCB

          ....|....| ....|....| ....|....| ....|....| ....|....|
            155        165        175        185        195
B lic     AAAAAAGCTG ACCGCATCGC GGCTGAAGGC TTGACTTTGA TCAAAACCGA
B hal     AAAAAAGCGG ACCGCGTAGC TGCTGAAGGA CTAGCCTATG TAAAAGCAGA
B sub     AAAAAAGCTG ACCGTATCGC AGCAGAAGGT TCTACTCTTA TTAAAACTGA
B cer     AAAAAAGCAG ACCGCATCGC TGCTGAAGGT TTAACTTTCA TCGAAACAAA
B ant     AAAAAAGCAG ACCGCATCGC TGCTGAAGGT TTAACTTTCA TCGAAACAAA
S aur     AAAAAAGCAG ACCGTATTGC GGCTGAAGGT TTAGTACATG TAGAAACTAA
L mon     AAAAAATCTG ATCGTGTTGC TTCTGAAGGT ATGACTCATG TAATCAGCAA
O ihe     AAAAAAGCAG ATCGTGTAGC AGCAGAAGGA TTAACTCATA TTGAGGTTGA
Conse     AAAAAAKCDG AYCGYRTHGC DKCWGAAGGH HYDRYHYWBR THRWVRBHRA
```

Figure 5.1. BLAST results for *tsf* genes chosen for degenerate primers. Forward (starting from nt 45) and reverse (starting from nt 679) primer sites shown as bold.

B lic: *Bacillus licheniformis*, B hal: *Bacillus halodurans*, B sub: *Bacillus subtilis*, B cer: *Bacillus cereus*, B ant: *Bacillus anthrasis*, S aur: *Staphylococcus aureus*, L mon: *Listeria monocytogenes*, O ihe: *Oceanobacillus iheyensis*, Conse: Consensus

```
          ....|....| ....|....| ....|....| ....|....| ....|....|
            205        215        225        235        245
B lic   CGGCAACACA GGCGTTATTC TTGAAGTGAA CTCTGAAACG GACTTCGTTG
B hal   AGGGAATCAC GCGATCATCG TCGAAGTTAA CTCTGAAACA GACTTCGTCG
B sub   CGGCAACAAA GGCGTTATCC TAGAAGTAAA CTCTGAAACT GATTTCGTTG
B cer   CGGTAACGAC GGTTTAATCT TAGAATTAAA CTCTGAAACT GATTTCGTTG
B ant   CGGTAACGAC GGTTTAATCT TAGAATTAAA CTCTGAAACT GATTTCGTTG
S aur   AGGTAACGAC GCAGTTATCG TTGAAATCAA CTCTGAAACA GACTTTGTTG
L mon   TGAAAAACAT GCCGTAGTAC TTGAAGTAAA TGCTGAAACA GATTTCGTTG
O ihe   AGGTAATAAA GCTGCAATAA TTGAAGTTAA TTGTGAAACT GACTTTGTTA
Conse   HGRNAAHVMH GSNDYHRTHN THGAADTNAA YKSTGAAACD GAYTTYGTYR

          ....|....| ....|....| ....|....| ....|....| ....|....|
            255        265        275        285        295
B lic   CGAAAAACGA AGGCTTCAAG GAGCTTTTGA ACGACCTGGC TGACCACATT
B hal   CAAAAAACGA AAACTTCCAA AAACTCGTAG CTGAATTAGC TTCTCACTTG
B sub   CGAAAAACGA AGGTTTTAAA GAGCTTCTTA ACACTTTAGC TGACCACCTT
B cer   CGAAAAACGA AGGTTTCCAA ACATTAATTA AAGAATTAGC TGCTCACTTA
B ant   CGAAAAACGA AGGTTTCCAA ACATTAATTA AAGAATTAGC TGCTCACTTA
S aur   CTCGTAACGA AGGATTCCAA GAGTTAGTTA AAGAAATCGC TAATCAAGTA
L mon   CTAAAAACGA TAACTTCCAA CAATTAGTTG ACGCTTTAGC TAAACAAATT
O ihe   CAAAAAACGA TCAGTTCAAA CAATTGTTAA GTGAGCTTGG TAAGCATATT
Conse   CDMRWAACGA WVRNTTYMAR VMRYTNNTDR VHRMNHTNGS TDMNCAHNTD

          ....|....| ....|....| ....|....| ....|....| ....|....|
            305        315        325        335        345
B lic   CTGGCTGAAA AGCCTGAAAG CGTTGAAGCT GCTATGGGTC AAAAAATGGC
B hal   CTAGAAAAAC GTCCTGCTTC TGTGGAAGAA GCGCTAGAGC AACCATTTAA
B sub   CTTGCAAACA CGCCAGCTGA TGTTGAAGAA GCAATGGGCC AAAAAATGGA
B cer   TTAGCTAACA AACCAGCTAA CGTTGAAGAA GCTATGGCTC AAACAATGGA
B ant   TTAGCTAACA AACCAGCTAA CGTTGAAGAA GCTATGGCTC AAACAATGGA
S aur   TTAGATACAA AAGCTGAAAC TGTTGAAGCT TTAATGGAAA CAACTTTACC
L mon   CTTGCAGTTC GTCCAGATAG CTTAGAAGAT GCGCTTAAAA CAGAAATGCC
O ihe   GTTGCAAATG AACCAGCGAC AGTTGAAGAA GCTCTTCAAC AAAAACTTCA
Conse   BTDGMWRHHV VDSCWGMDDV HKTDGAAGMW KYDMTDVVNM MAVMWHTDVM

          ....|....| ....|....| ....|....| ....|....| ....|....|
            355        365        375        385        395
B lic   TA---ACGGT TCAACTGTAG AAGAATACAT CACAAGCGCG GTTGCTAAAA
B hal   CG---GTGGA GAAACGGTTC AAGAATACAT CAACTCTGCG ATTGCAAAAA
B sub   AA---ATGGC TCAACTGTTG AAGAGTACAT CACAAGCGCA GTTGCTAAAA
B cer   AA---ACGGC AAAAAAGTAG AAGAGCACAT CAACGAAGCA ATCGCTAAAA
B ant   AA---ACGGC AAAAAAGTAG AAGAGCACAT CAACGAAGCA ATCGCTAAAA
S aur   AA---ATGGT AAATCAGTTG ATGAAAGAAT TAAAGAAGCA ATTTCAACAA
L mon   TA---ATGGT CAAACTGTTC AAGACTACAT CACTGAAGCA ATTACAAAAA
O ihe   TGGAGATGGA GAAACAGTTG AATCTGTAAT CACAAACGCA GTTGCAAAAA
Conse   HRGAGRYGGH NMAWMDGTWS AWKMNNDMAT YAMHDVHGCR RTYDCWAMAA
```

Figure 5.1. BLAST results for *tsf* genes chosen for degenerate primers, forward and reverse primer sites shown as bold, forward primer site starting from 45, reverse primer site starting from 679

B lic: *Bacillus licheniformis*, B hal: *Bacillus halodurans*, B sub: *Bacillus subtilis*, B cer: *Bacillus cereus*, B ant: *Bacillus anthrasis*, S aur: *Staphylococcus aureus*, L mon: *Listeria monocytogenes*, O ihe: *Oceanobacillus iheyensis*, Conse: Consensus, (continued)

```
          ....|....| ....|....| ....|....| ....|....| ....|....|
             405        415        425        435        445
B lic     TCGGAGAAAA AATCACGCTT CGCCGCTTCG CTGTCCTGAC AAAAGGCGAT
B hal     TTGGCGAAAA GCTTTCCCTT CGTCGGTTTG AAATTGTTGA AAAAGAAGAC
B sub     TCGGTGAGAA AATCACTCTT CGCCGCTTTA CAGTTCTTAC AAAAGACGAC
B cer     TTGGTGAAAA ACTTACACTT CGTCGTTTCG AAATCGTATC AAAAACTGAT
B ant     TTGGTGAAAA ACTTACACTT CGTCGTTTCG AAATCGTATC AAAAACTGAT
S aur     TCGGTGAAAA ATTAAGTGTT CGTCGTTTTG CTATCAGAAC TAAAACTGAT
L mon     TCGGTGAAAA CATTTCCCTT CGTCGTTTTG AAGTAAAAGA AAAAGCAGAC
O ihe     TTGGTGAGAA AATTTCCTTA CGTCGTTTTG AAGTACTAGA AAAAACTGAT
Conse     TYGGHGARAA VHTHWSNBTW CGYCGBTTYR MWRTHVDDDM WAAARVHGAY

          ....|....| ....|....| ....|....| ....|....| ....|....|
             455        465        475        485        495
B lic     GACGCTGCAT TCGGCGCATA CCTGCACATG GGCGGAAAAA TCGGCGTATT
B hal     GGAGATGTGT TTGGACAGTA CATCCATATG GGTGGACGCA TCGGGGTTCT
B sub     AGCTCTGCAT TCGGTGCGTA CCTTCACATG GGCGGCCGCA TCGGTGTATT
B cer     GCAGATGCAT TCGGCGCTTA CCTACACATG GGTGGACGCA TTGGTGTACT
B ant     GCAGATGCAT TCGGCGCTTA CCTACACATG GGTGGACGCA TTGGTGTATT
S aur     AACGATGCTT TCGGCGCTTA CTTACACATG GGTGGACGCA TTGGTGTATT
L mon     AACTCTGCTT TCGGTGAATA CATCCACATG AACGGACGTA TTGGTGTTCT
O ihe     AACGATGCAT TTGGTGCTTA CTTGCACATG GGTGGTACTA TTGGTGTACT
Conse     RVMKMTGYDT TYGGHSMDTA CHTNCAYATG RRYGGHMVHA TYGGBGTWYT

          ....|....| ....|....| ....|....| ....|....| ....|....|
             505        515        525        535        545
B lic     GACTGTCCTT AACGGCACAA CTGACGAAGA AACAGCAAGA GACATCGCAA
B hal     TTCCGTTATC GGTCAATCTT CAGATGAAGA GCTTGCGAAG GATATTGCGA
B sub     AACTGTTCTC AACGGTACAA CTGATGAAGA AACTGCGAAA GATATCGCAA
B cer     AACAGTTCTT GAAGGTTCTA CTGATGAAGC GGCTGCTAAA GATGTTGCAA
B ant     AACAGTTCTT GAAGGATCTA CTGATGAAGC GGCTGCTAAA GATGTTGCAA
S aur     AACAGTTGTT GAAGGTTCAA CTGACGAAGA AGCAGCAAGA GACGTTGCTA
L mon     TACACTTCTT GAAGGAACTA CTGATACTAC TGTTGCAAAA GACGTTGCAA
O ihe     ATCATTGTTA GAAGGTACTA CAGATGAACA AGTTGGAAAA GATATCGCTA
Conse     DWCHBTBNTH RRHSRHWCWW CWGAYRMWVM DVYWGSDARR GAYRTYGCDA

          ....|....| ....|....| ....|....| ....|....| ....|....|
             555        565        575        585        595
B lic     TGCACGTTGC AGCTGTAAAC CCTCGCTTCA TTTCCCGCGA CCAAGTGTCA
B hal     TGCACGTTGC AGCGATTAAC CCTACTTATG TGACACGCGA TCAAGTGTCT
B sub     TGCACGTTGC TGCAGTTAAC CCTCGTTACA TTTCTCGTGA TCAAGTATCT
B cer     TGCACATCGC TGCAGTTAAC CCTAAATACA TCGACCGCGA TGCTGTAACA
B ant     TGCACATCGC TGCAGTTAAC CCTAAATACA TCGACCGCGA TGCTGTAACA
S aur     TGCATATCGC TGCAATCAAC CCTAAATATG TTTCTTCTGA ACAAGTTAGC
L mon     TGCACATCGC TGCAATCAAC CCTAAATACA TTTCTCGTGA AGACGTTTCT
O ihe     TGCACGTTGC AGCAGTAAAT CCTCGTTATG TAACTCGTGA TGAAGTTGCT
Conse     TGCAYRTYGC WGCDRTHAAY CCTMVHTWYR TNDMHYSYGA HSMHGTDDSH
```

Figure 5.1. BLAST results for *tsf* genes chosen for degenerate primers, forward and reverse primer sites shown as bold, forward primer site starting from 45, reverse primer site starting from 679

B lic: *Bacillus licheniformis*, B hal: *Bacillus halodurans*, B sub: *Bacillus subtilis*, B cer: *Bacillus cereus*, B ant: *Bacillus anthrasis*, S aur: *Staphylococcus aureus*, L mon: *Listeria monocytogenes*, O ihe: *Oceanobacillus iheyensis*, Conse: Consensus, (continued)

```
          ....|....| ....|....| ....|....| ....|....| ....|....|
             605        615        625        635        645
B lic     GAAGAAGAAG CGAACCGCGA GCGCGAAATC TTGACTCAGC AGGCGCTTCA
B hal     GAGGATGAAG TAGCTCGCGA GCGCGAAGTG TTAAAACAGC AAGCACTTAA
B sub     GAAGAAGAAA CAAATCATGA GCGTCAAATC TTAACTCAGC AAGCCCTCCA
B cer     GCTGAAGAAG TTGAGCATGA GCGTCAAGTA TTAACACAAC AAGCATTAAA
B ant     GCTGAAGAAG TTGAGCATGA GCGTCAAGTA TTAACACAAC AAGCATTAAA
S aur     GAAGAAGAAA TCAACCACGA AAGAGAAGTT TTAAAACAAC AAGCATTAAA
L mon     ACTGAAGAAG TGGAACACGA AAAAGAAGTA TTAACTCAAC AAGCGTTAAA
O ihe     GAAGAAGAAG TGAATCGTGA ACGTGAAGTG TTGAAAACAC AAGCATTAAA
Conse     RMDGAWGAAR YNRMNCRYGA RMRHSAARTN TTRAMWMMRC ARGCVYTHMA

          ....|....| ....|....| ....|....| ....|....| ....|....|
             655        665        675        685        695
B lic     AGAAGGCAAC CCTGAAAACA TCGTCGCA**AA AATGGTTGAA GGCCGC**CTGA
B hal     TGAAGGTAAG CCAGAAAACA TCGTCGAA**AA GATGGTGGAA GGGCGT**CTTG
B sub     AGAAGGCAAA CCTGAAAACA TCGTAGCG**AA AATGGTTGAA GGCCGT**CTGA
B cer     CGAAGGCAAG CCTGAAAAAA TCGTTGCA**AA AATGGTTGAA GGCCGT**CTTG
B ant     CGAAGGCAAG CCTGAAAAAA TCGTTGCA**AA AATGGTTGAA GGCCGT**CTTG
S aur     TGAAGGTAAA CCAGAAAACA TCGTTGAA**AA AATGGTGGAA GGACGT**TTAC
L mon     CGAAGGCAAA CCAGCTAATA TCGTTGAA**AA AATGGTAGAA GGCCGT**TTGA
O ihe     TGAAGGTAAA CCAGAGAATA TTGTTGAG**AA AATGGTTGAA GGTCGT**CTTG
Conse     HGAAGGYAAV CCWGMDAAHA TYGTHGMR**AA RATGGTDGAA GGNCGY**YTDV

          ....|....| ....|....| ....|....| ....|....| ....|....|
             705        715        725        735        745
B lic     ACAAATATTT CGAAGAAATC TGCCTGCTTG ATCAAGCGTT CGTTAAAAAC
B hal     GCAAATACTT TGAGCAAGTA TGTTTACTTG ACCAAGCATT CGTTAAAGAT
B sub     ACAAATTCTT CGAAGAAATT TGTCTATTAG ACCAAGCGTT CGTTAAAAAC
B cer     GCAAATTCTT CGAAGAAATT TGCTTACTTG ACCAAGCATT CGTTAAAAAC
B ant     GCAAATTCTT CGAAGAAATT TGCTTACTTG ACCAAGCATT CGTTAAAAAC
S aur     GTAAATACTT ACAAGAAATT TGTGCTGTAG ATCAAGACTT CGTTAAAAAC
L mon     AAAAATATCT AAGCGAAATT TCCTTAGAAG ACCAACCTTT CGTTAAAAAC
O ihe     GTAAATTCTT TGAAGACATC GTATTATTAG AACAAAGCTT CGTTAAAGAT
Conse     RHAAATWYYT HVRVSAMRTH KBHBYDBWWG AHCAAVVNTT CGTTAAARAY

          ....|....| ....|....| ....|....| ....|....| ....|....|
             755        765        775        785        795
B lic     CCGGATGAAA AAGTGAAACA GGTCGTTGCT GCGAAAAACG CAACGGTTGA
B hal     GGCGATCAAA AAGTTGGCAA ATACGTGCAA AGCAAAGGGG CTACTGTGAA
B sub     CCAGATGAAA AAGTGAAACA AGTTATTGCA GCGAAAAACG CTACTGTTCA
B cer     CCTGATATGA AAGTTCGTCA GTTCGTTGAG TCTAAAGGCG GAACATTAAA
B ant     CCTGATATGA AAGTTCGTCA GTTCGTTGAG TCTAAAGGCG GAACATTAAA
S aur     CCTGATGTAA CAGTTGAAGC TTTCTTAAAA ACAAAAGGTG GAAAACTTGT
L mon     CCAGACATCA CTGTAGGTGA CTACGTGAAA CAAAGCGGTG GTAAAGTTGT
O ihe     CCAGACCAAA AAGTGAAAAA ATATGTAGCT GATAAAGGTG CGGCTGTAAA
Conse     SSNGAYVWVA MWGTDVRHVM NKWYDTDVMD NVNARMRRBG SDRMDBTDVW
```

Figure 5.1. BLAST results for *tsf* genes chosen for degenerate primers, forward and reverse primer sites shown as bold, forward primer site starting from 45, reverse primer site starting from 679

B lic: *Bacillus licheniformis*, B hal: *Bacillus halodurans*, B sub: *Bacillus subtilis*, B cer: *Bacillus cereus*, B ant: *Bacillus anthrasis*, S aur: *Staphylococcus aureus*, L mon: *Listeria monocytogenes*, O ihe: *Oceanobacillus iheyensis*, Conse: Consensus, (continued)

```
          ....|....| ....|....| ....|....| ....|....| ....|....|
             805        815        825        835        845
B lic     AACGTTCGTC CGCTACGAAG TTGGAGAAGG TATCGAAAAA CGCCAAGAAA
B hal     AGAGTTCATC CGCTATGAAG TAGGCGAAGG GCTCGAGAAG CGTGAAGACA
B sub     AACTTTTGTC CGCTATGAAG TTGGAGAAGG CATCGAAAAA CGTCAAGAAA
B cer     AGGATTCGTT CGCTACGCTG TTGGTGAAGG TATCGAAAAA CGCGAAGACA
B ant     AGGATTCGTT CGCTACGCTG TTGGTGAAGG TATCGAAAAA CGCGAAGACA
S aur     TGACTTCGTA CGCTATGAAG TAGGCGAAGG TATGGAAAAA CGCGAAGAAA
L mon     ATCATTCGTA CGTTTCGAAG TAGGCGAAGG AATCGAGAAA AAAGAAGATA
O ihe     AACATTTGTT CGTTATGAAG TAGGCGAAGG AATGGAAAAA CGCGAAGAAA
Conse     WDVNTTYRTH CGYTWYGMWG TWGGHGAAGG NMTSGARAAR MRHSAAGAHA

          ....|....| ....|....| ....|....| ....|....| ....|
             855        865        875        885        895
B lic     ACTTTGCTGA AGAAGTTATG AACCAAGTGA AAA--AATAA ----
B hal     ACTTTGCAGA AGAAGTTATG TCTCAAGTGA AGA--AGTAA ----
B sub     ACTTTGCTGA AGAAGTAATG AACCAAGTGA AAA--AATAA ----
B cer     ACTTTGCTGA AGAAGTAATG AACCAAGTAA AAGGTAGTAA CTAA
B ant     ACTTTGCTGA AGAAGTAATG AACCAAGTAA AAGGTAGTAA CTAA
S aur     ACTTTGCGGA TGAAGTTAAA GGACAAATGA AAT--AA--- ----
L mon     ACTTTGTTGA AGAAGTAATG AGCCAAGTGA AAA--AATAA ----
O ihe     ACTTCGCTGA AGAAGTAATG AGTCAAATTA AAA--AATAG ----
Conse     ACTTYGYDGA WGAAGTWAWR DVHCAARTDA ARDGTARTAR CTAA
```

Figure 5.1. BLAST results for *tsf* genes chosen for degenerate primers, forward and reverse primer sites shown as bold, forward primer site starting from 45, reverse primer site starting from 679

B lic: *Bacillus licheniformis*, B hal: *Bacillus halodurans*, B sub: *Bacillus subtilis*, B cer: *Bacillus cereus*, B ant: *Bacillus anthrasis*, S aur: *Staphylococcus aureus*, L mon: *Listeria monocytogenes*, O ihe: *Oceanobacillus iheyensis*, Conse: Consensus, (continued)

5.1.2. Extraction of Genomic DNA from *Geobacillus anatolicus*

Extraction of genomic DNA from *Geobacillus anatolicus* was done as described in section 4.1. Genomic DNA was visualized as a homogeneous band after electrophoresis on 1 % agarose gels (Figure 5.2).

5.1.3. Partial PCR Amplification of *Geobacillus anatolicus tsf* Gene for Primary Information About The *tsf* Gene

Degenerate primer pair TsF45/TsR696 was designed corresponding to the most conserved regions of the known *tsf* genes from phylogenetically related bacteria to *Geobacillus anatolicus*.

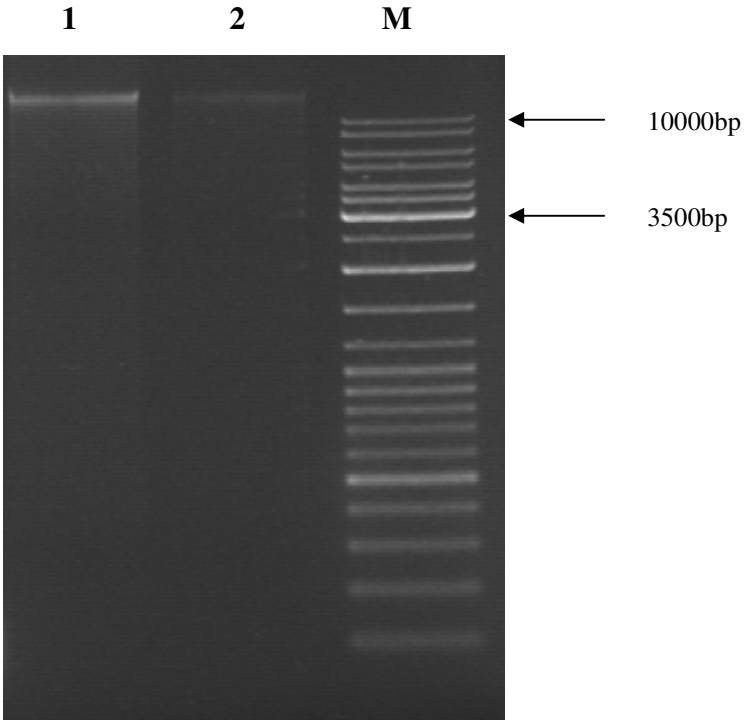

Figure 5.2. Genomic DNA from *Geobacillus anatolicus*. 1-2, Genomic DNA in two different concentrations; M, Molecular weight marker

Estimated product size after the PCR amplification using TsF45/TsR696 primer pair was about 651 bp, as calculated from *Bacillus halodurans C-125* and *Bacillus subtilis subsp. subtilis str. 168 tsf* sequences.

PCR amplifications using annealing temperature gradient between 50.0^{o}C to 60.0^{o}C gave a unique band for each case when analyzed by electrophoresis on 1 % agarose gels (Figure 5.3).

5.1.4. Confirmation of *Geobacillus anatolicus tsf* Gene Sequence

The PCR product obtained from partial amplification of *Geobacillus anatolicus* elongation factor Ts gene was sequenced using TsF45 and TsR696 primers A partial 500 bp-long sequence was obtained (data not shown). The sequencing results were analyzed by aligning with the known elongation factor Ts gene sequences confirming that the PCR amplified fragment was in fact part of the *tsf* gene.

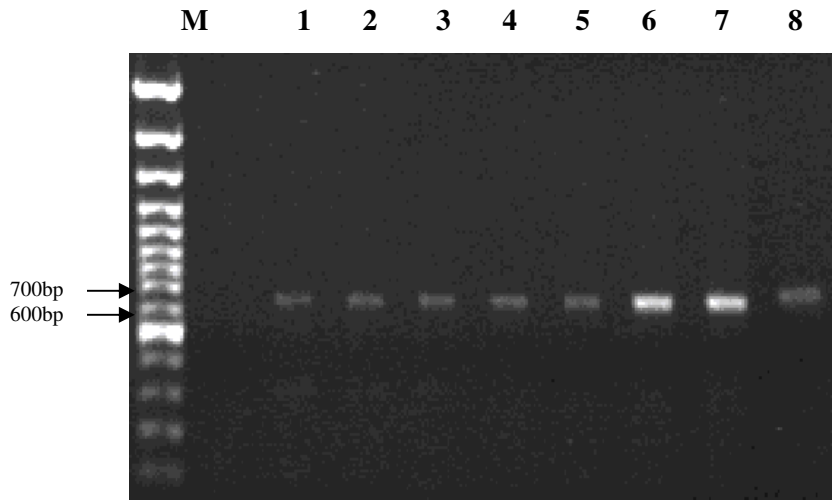

Figure 5.3. PCR amplification of *Geobacillus anatolicus tsf* gene with TsF45 and TsR696 primer pair
M, molecular size marker; 1-8, annealing temperature gradient (50, 50.7, 52, 53.7, 56.2, 58.1, 59.3, 60°C respectively).

5.2. Complete Sequence Data for *Geobacillus anatolicus tsf* Gene

5.2.1. Constructing Degenerate Primers for PCR Fragment Containing Complete *tsf* Gene

Complete genomic data is available for some bacteria closely related to *Geobacillus anatolicus*. In these organisms *tsf* gene lies between *rpsB* and *smbA* genes, coding for 30S ribosomal protein S2 and uridylate kinase genes, respectively (see Figure 1.5 and Figure 3.1). *rpsB* and *smbA* gene sequences of the organisms given in Table 5.2 were used for designing degenerate primers.

Table 5.2. Bacterial *rpsB* genes used for primer designation

Source organism	**Gene ID**	**Nucleic acid sequence length**
Bacillus licheniformis ATCC 14580	3031061	741 bp
Bacillus halodurans C-125	890719	735 bp
Bacillus subtilis subsp. subtilis str. 168	939599	741 bp
Bacillus cereus E33L	3022973	702 bp
Bacillus anthracis str. 'Ames Ancestor'	2820094	702 bp
Bacillus thuringiensis serovar konkukian str. 97-27	2858304	702 bp

Table 5.3. Bacterial *smbA* genes used for primer designation

Source organism	**Gene ID**	**Nucleic acid sequence length**
Geobacillus kaustophilus HTA426	3185885	723 bp
Bacillus thuringiensis serovar konkukian str. 97-27	2856144	723 bp
Bacillus cereus E33L	3025654	723 bp
Bacillus halodurans C-125	890720	720 bp
Bacillus subtilis subsp. subtilis str. 168	939602	723 bp
Oceanobacillus iheyensis HTE831	1017976	723 bp

```
          ....|....| ....|....| ....|....| ....|....| ....|....|
              5         15         25         35         45
B sub     ATGTCAGTCA TTTCTATGAA GCAATTGCTT GAAGCTGGTG TTCACTTCGG
B thu     ATGTCAGTAA TTTCTATGAA GCAATTGCTT GAAGCTGGTG TTCATTTCGG
B lic     ATGTCAGTCA TTTCTATGAA GCAATTGCTT GAAGCTGGTG TTCACTTCGG
B hal     GTGGCAGTTA TTTCCATGAA ACAATTGTTG GAAGCTGGGG TGCACTTCGG
B cer     ATGTCAGTAA TTTCTATGAA GCAATTGCTT GAAGCTGGTG TTCATTTTGG
B ant     ATGTCAGTAA TTTCTATGAA GCAATTGCTT GAAGCTGGTG TTCATTTCGG
Conse     RTGKCAGTHA TTTCYATGAA RCAATTGYTK GAAGCTGGKG TKCAYTTYGG

          ....|....| ....|....| ....|....| ....|....| ....|....|
              55         65         75         85         95
B sub     TCACCAAACA CGCCGTTGGA ACCCAAAAAT GAAGCGTTAC ATTTTCACGG
B thu     ACATCAAACT CGTCGTTGGA ACCCAAAAAT GAAGCGTTAC ATTTTCACAG
B lic     CCACCAAACG CGCCGCTGGA ACCCAAAAAT GAAGCGTTAC ATCTTCACGG
B hal     TCACCAAACA CGCCGTTGGA ACCCTAAAAT GGATCGCTAC ATCTTCACAG
B cer     ACATCAAACT CGTCGTTGGA ACCCAAAAAT GAAGCGTTAC ATTTTCACAG
B ant     ACATCAAACT CGTCGTTGGA ACCCAAAAAT GAAGCGTTAC ATTTTCACAG
Conse     HCAYCAAACD CGYCGYTGGA ACCCWAAAAT GRAKCGYTAC ATYTTCACRG

          ....|....| ....|....| ....|....| ....|....| ....|....|
             105        115        125        135        145
B sub     AGCGTAACGG AATCTACATC ATTGACCTTC AAAAAACAGT CAAAAAAGTA
B thu     AGCGTAACGG TATCTACATC ATCGACTTAC AAAAAACTGT GAAAAAAGTT
B lic     AGCGTAACGG AATCTACATC ATCGATCTTC AAAAAACAGT CAAAAAAGTT
B hal     AACGTAACGG AATTTACATT ATCGATTTAC AAAAGACCGT CAAAAAAGTA
B cer     AGCGTAACGG TATCTACATC ATCGACTTAC AAAAAACTGT GAAAAAAGTT
B ant     AGCGTAACGG TATCTACATC ATCGACTTAC AAAAAACTGT GAAAAAAGTT
Conse     ARCGTAACGG WATYTACATY ATYGAYYTWC AAAARACHGT SAAAAAAGTW

          ....|....| ....|....| ....|....| ....|....| ....|....|
             155        165        175        185        195
B sub     GAGGAAGCTT ACAACTTCAC TAAAAATCTT GCTGCTGAAG GCGGAAAAAT
B thu     GAGGAAGCTT TCAAAGTTAT GCGTGACATC GCTGCTGAAG GCGGAGACAT
B lic     GAAGAAGCAT ACAACTTCAC GAAAAACCTT GCAGCTGACG GAGGAAAAAT
B hal     GAGGAAGCGT ACAATTTTGT TCGCGAACTT GCTGCTGACG GCGGGAAAGT
B cer     GAGGAAGCTT TCAAAGTTAT GCGTGACATC GCTGCTGAAG GCGGAGACAT
B ant     GAGGAAGCTT TCAAAGTTAT GCGTGACATC GCTGCTGAAG GCGGAGACAT
Conse     GARGAAGCDT WCAAHKTYRY KMRHRAHMTY GCWGCTGAMG GMGGRRAMRT

          ....|....| ....|....| ....|....| ....|....| ....|....|
             205        215        225        235        245
B sub     CCTTTTCGTC GGAACAAAAA AACAAGCTCA AGATTCTGTT AAAGAAGAAG
B thu     CTTATTCGTA GGTACTAAAA AACAAGCACA AGAAGCTATT AAAGAAGAAG
B lic     CCTTTTCGTC GGTACGAAAA AGCAAGCTCA AGACTCTGTT AAAGAAGAAG
B hal     CCTTTTCGTA GGGACGAAAA AGCAAGCGCA AGATTCAGTG AAGGAAGAAG
B cer     CTTATTCGTA GGTACTAAAA AACAAGCACA AGAAGCTATT AAAGAAGAAG
B ant     CTTATTCGTA GGTACTAAAA AACAAGCACA AGAAGCTATT AAAGAAGAAG
Conse     CYTWTTCGTM GGDACDAAAA ARCAAGCDCA AGAHKCWRTK AARGAAGAAG
```

Figure 5.4. BLAST results for *rpsB* (upstream to *tsf*) gene used for forward primer design (Primer site is indicated in bold characters)

B sub: *Bacillus subtilis*, B thu: *Bacillus thuringiensis serovar konkukian*, B lic: *Bacillus licheniformis*, B hal: *Bacillus halodurans*, B cer: *Bacillus cereus*, B ant: *Bacillus anthrasis*, Conse: Consensus

```
          ....|....| ....|....| ....|....| ....|....| ....|....|
             255        265        275        285        295
B sub     CTCAACGCTC TGGCATGTAC TATGTCAACC AACGCTGGTT GGGCGGTACA
B thu     CAACTCGTGC TGGTATGTAC TTCGTTAACC AACGTTGGTT AGGTGGAACT
B lic     CGGAACGTTC TGGAATGTAC TATGTCAACC AACGCTGGCT TGGCGGTACA
B hal     CTGAGCGCTG CGGTATGTAC TATATCAACC AACGTTGGTT AGGTGGTACA
B cer     CAACTCGTGC TGGTATGTAC TTCGTTAACC AACGTTGGTT AGGTGGAACT
B ant     CAACTCGTGC TGGTATGTAC TTCGTTAACC AACGTTGGTT AGGTGGAACT
Conse     CDVMDCGYKS YGGHATGTAC TWYRTYAACC AACGYTGGYT DGGYGGWACW

          ....|....| ....|....| ....|....| ....|....| ....|....|
             305        315        325        335        345
B sub     TTAACAAACT TCGAAACAAT CCAAAAACGT ATTAAACGTC TTAAAGACAT
B thu     TTAACTAACT TCCAAACAAT CCAAAAGCGT ATCAAGCGTC TTAAAGACAT
B lic     TTAACAAACT TTGAAACGAT CCAAAAGCGT ATCAAACGCC TTAAAGATAT
B hal     TTAACGAACT TTGAGACCAT TCAAAAGCGC ATTGATCGCC TTAAAAAGCT
B cer     TTAACTAACT TCCAAACAAT CCAAAAGCGT ATCAAGCGTC TTAAAGACAT
B ant     TTAACTAACT TCCAAACAAT CCAAAAGCGT ATCAAGCGTC TTAAAGACAT
Conse     TTAACDAACT TYSARACVAT YCAAAARCGY ATYRADCGYC TTAAARABMT

          ....|....| ....|....| ....|....| ....|....| ....|....|
             355        365        375        385        395
B sub     TGAAAAAATG CAAGAAAACG GTACGTTTGA TGTACTTCCT AAAAAAGAAG
B thu     CGAAAGAATG CAAGAAGATG GTACTTTCGA AGTACTACCT AAGAAAGAAG
B lic     CGAAAAAATG CAGGAAAACG GCACATTTGA CGTACTTCCT AAAAAAGAAG
B hal     CGAGAAAATG GAAGAGGATG GAACGTTCGA TGTTCTTCCT AAGAAAGAAG
B cer     CGAAAGAATG CAAGAAGATG GTACTTTCGA AGTACTACCT AAGAAAGAAG
B ant     CGAAAGAATG CAAGAAGATG GTACTTTCGA AGTACTACCT AAGAAAGAAG
Conse     YGARARAATG SARGARRAYG GHACDTTYGA HGTWCTWCCT AARAAAGAAG

          ....|....| ....|....| ....|....| ....|....| ....|....|
             405        415        425        435        445
B sub     TCGTTCAATT GAAAAAAGAA TTAGAGCGTC TTGAAAAATT CCTAGGCGGA
B thu     TTGTTCAACT TAAAAAAGAG TTAGAGCGTC TTGAGAAATT CTTAGGCGGT
B lic     TCGTTCAATT GAAAAAAGAA TTAGAGCGTC TTGAAAAATT CCTCGGCGGA
B hal     TTATCCTTCT TAAGAAAGAG ATGGAGCGTC TTGAGAAGTT CCTAGGCGGA
B cer     TTGTTCAACT TAAAAAAGAG TTAGAGCGTC TTGAGAAATT CTTAGGCGGT
B ant     TTGTTCAACT TAAAAAAGAG TTAGAGCGTC TTGAGAAATT CTTAGGCGGT
Conse     TYRTYCWWYT KAARAAAGAR WTRGAGCGTC TTGARAARTT CYTMGGCGGW

          ....|....| ....|....| ....|....| ....|....| ....|....|
             455        465        475        485        495
B sub     ATTAAAGATA TGAAGGATCT TCCTGATGCA TTATTCATCA TCGATCCTCG
B thu     ATTAAAGATA TGAAAGGTCT TCCAAGTGCA TTATTCGTAG TAGACCCTCG
B lic     ATCAAAGAAA TGAAAGAGCT TCCTGACGCA CTGTTCATCA TCGATCCTCG
B hal     ATTAAAGATA TGAACAGCCT TCCGGATGCC CTTTTTGTGA TCGACCCTCG
B cer     ATTAAAGATA TGAAAGGTCT TCCAAGTGCA TTATTCGTAG TAGACCCTCG
B ant     ATTAAAGATA TGAAAGGTCT TCCAAGTGCA TTATTCGTAG TAGACCCTCG
Conse     ATYAAAGAWA TGAAVRRBCT TCCDRRYGCM YTDTTYRTVR TMGAYCCTCG
```

Figure 5.4. BLAST results for *rpsB* (upstream to *tsf*) gene used for forward primer design (Primer site is indicated in bold characters)

B sub: *Bacillus subtilis*, B thu: *Bacillus thuringiensis serovar konkukian*, B lic: *Bacillus licheniformis*, B hal: *Bacillus halodurans*, B cer: *Bacillus cereus*, B ant: *Bacillus anthrasis*, Conse: Consensus (continued)

```
          ....|....| ....|....| ....|....| ....|....| ....|....|
             505        515        525        535        545
B sub     CAAAGAGCGC ATCGCTGTTG CGGAAGCTCG CAAATTAAAC ATCCCTATCA
B thu     TAAAGAGCGT ATTGCAGTTG CTGAAGCACG CAAATTACAC ATTCCAATCA
B lic     CAAAGAGCGC ATCGCGGTTG CAGAAGCGCG CAAACTGAAC ATTCCGATCA
B hal     TAAAGAGCGT ATTGCGATCG CGGAGGCTCA TAAGCTAAAC ATTCCAATCG
B cer     TAAAGAGCGT ATTGCAGTTG CTGAAGCACG CAAATTACAC ATTCCAATCA
B ant     TAAAGAGCGT ATTGCAGTTG CTGAAGCACG CAAATTACAC ATTCCAATCA
Conse     YAAAGAGCGY ATYGCDRTYG CDGARGCDCR YAARYTRMAC ATYCCDATCR

          ....|....| ....|....| ....|....| ....|....| ....|....|
             555        565        575        585        595
B sub     TCGGTATCGT AGATACTAAC TGTGATCCAG ATGAAATCGA TGTTGTTATC
B thu     TCGGTATCGT TGATACAAAC TGTGATCCAG ACGAAATCGA TCACGTTATC
B lic     TCGGAATCGT TGACACAAAC TGCGATCCTG ATGAAATCGA TGTTGTCATC
B hal     TAGCGATTGT AGATACAAAC TGTGACCCAG ATGAAATCGA TTATGTCATT
B cer     TCGGTATCGT TGATACAAAC TGTGATCCAG ACGAAATCGA TCACGTTATC
B ant     TCGGTATCGT TGATACAAAC TGTGATCCAG ACGAAATCGA TCACGTTATC
Conse     TMGSDATYGT WGAYACWAAC TGYGAYCCWG AYGAAATCGA TBWYGTYATY

          ....|....| ....|....| ....|....| ....|....| ....|....|
             605        615        625        635        645
B sub     CCAGCGAACG ATGACGCTAT CCGCGCTGTT AAACTTCTAA CTTCTAAAAT
B thu     CCAGCAAACG ATGATGCAAT TCGTGCTGTA AAACTTCTTA CATCTAAAAT
B lic     CCTGCGAACG ATGACGCCAT CCGCGCTGTT AAATTGCTGA CTTCTAAAAT
B hal     CCAGGTAACG ATGACGCGAT TCGCGCGGTT AAACTTTTGA CTGGAAAAAT
B cer     CCAGCAAACG ATGATGCAAT TCGTGCTGTA AAACTTCTTA CATCTAAAAT
B ant     CCAGCAAACG ATGATGCAAT TCGTGCTGTA AAACTTCTTA CATCTAAAAT
Conse     CCWGSDAACG ATGAYGCNAT YCGYGCKGTW AAAYTKYTDA CWKSWAAAAT

          ....|....| ....|....| ....|....| ....|....| ....|....|
             655        665        675        685        695
B sub     GGCAGATGCA ATCTTAGAAG CGAAGCAAGG CGAAGAAGAA ------GCGG
B thu     GGCAGACGCG ATCCTTGAAG CAAAACAAGG TGAAGAAACT ----------
B lic     GGCAGATGCC ATCCTTGAAG CAAAACAAGG TGAAGAATCA ------GCTG
B hal     GGCAGATGCT GTAGTAGAAG CAACTTCTGG TG------CT ------GGCG
B cer     GGCAGACGCG ATCCTTGAAG CAAAACAAGG TGAAGAAACT ----------
B ant     GGCAGACGCG ATCCTTGAAG CAAAACAAGG TGAAGAAACT ----------
Conse     GGCAGAYGCN RTMBTWGAAG CRAMDYMWGG YGAAGAADMW ------GSBG

          ....|....| ....|....| ....|....| ....|....| ....|..
             705        715        725        735        745
B sub     AAGTTGCTGA AGAAACTGCA CCAGAAACAG AAACAACAAC TGCGTAA
B thu     --GTTACTGC GTAA------ ---------- ---------- -------
B lic     AAACAGAAGC GAAAGAAGCA GAAACAACTG AAACAACAAC TGCTTAA
B hal     AAGAGGCAGA AGAAGTGGCG GAAGTGACAG AAGAGACAAC AGCTTAA
B cer     --GTTACTGC GTAA------ ---------- ---------- -------
B ant     --GTTACTGC GTAA------ ---------- ---------- -------
Conse     AARHDRMWGM RDAARHDGCR SMARHRACWG AARMRACAAC WGCKTAA
```

Figure 5.4. BLAST results for *rpsB* (upstream to *tsf*) gene used for forward primer design (Primer site is indicated in bold characters)

B sub: *Bacillus subtilis*, B thu: *Bacillus thuringiensis serovar konkukian*, B lic: *Bacillus licheniformis*, B hal: *Bacillus halodurans*, B cer: *Bacillus cereus*, B ant: *Bacillus anthrasis*, Conse: Consensus (continued)

```
          ....|....| ....|....| ....|....| ....|....| ....|....|
              5         15         25         35         45
G kau     ATGGAACAGC CAAAATACAA ACGCGTCGTG TTAAAGTTGA GCGGGGAAGC
B thu     ATGAGTAAAC CGAAATATAA TCGTGTCGTT TTAAAGCTGA GTGGAGAAGC
B cer     ATGAGTAAAC CGAAATATAA TCGTGTCGTT TTAAAGCTGA GTGGAGAAGC
B hal     -----ATGGC TAAA-TATAA ACGAGTCGTT TTAAAGTTGA GCGGCGAGGC
O ihe     ATGACGACAG CACGTTACAA TAGAGTTGTG TTAAAATTAA GTGGAGAAGC
B sub     ATGGAAAAAC CAAAATACAA ACGTATCGTA TTAAAGCTTA GCGGGGAAGC
Conse     ATGRVDHVRS YRMRWTAYAA WMGHRTYGTD TTAAARYTDA GYGGVGARGC

          ....|....| ....|....| ....|....| ....|....| ....|....|
              55         65         75         85         95
G kau     GCTCGCCGGA AAGCAAGGGT TTGGCATTCA ACCGGCGGTC ATTCAGTCGA
B thu     TTTAGCTGGC GAGCAAGGAT TTGGAATTAA CCCAGCTGTT ATTAAATCAG
B cer     TTTAGCTGGC GAGCAAGGAT TTGGAATTAA CCCAGCTGTT ATTAAATCAG
B hal     GTTAGCAGGA GAACAAGGAT ATGGCATCGA TCCGGAAGTC ATTCAATCGA
O ihe     TTTAAGTGGT GATCAAGGAT ATGGAATCGA CCCAAAAGTA ATCCAATCCA
B sub     ATTGGCAGGA GAACAGGGAA ATGGAATTAA CCCGACTGTC ATTCAATCCA
Conse     DYTVRSHGGH RADCARGGRW WTGGMATYVA HCCRRMDGTH ATYMARTCVR

          ....|....| ....|....| ....|....| ....|....| ....|....|
             105        115        125        135        145
G kau     TCGCCAAACA AGTGAAAGAA GTTGTTGAAC TGGGCGTCGA GGTCGCCATC
B thu     TTGCGGAACA AGTGAAAGAA ATTGCAGAAC TTGATGTAGA GGTTGCTGTT
B cer     TTGCGGAACA AGTGAAAGAA ATTGCAGAAC TTGATGTAGA GGTTGCTGTT
B hal     TTGCTTCACA AATTAAAGAA ATCGTTGAAT TAGATGTAGA GGTTGCTGTC
O ihe     TTTCAAAGCA AGTAAAAGAG GTTGCTGACC TTGGTGTTGA AGTAGCCATT
B sub     TTGCAAAGCA AGTGAAGGAA ATCGCTGAGC TTGAAGTCGA AGTGGCTGTT
Conse     TYKCNDMRCA ARTDAARGAR RTYGYWGAVY TDGRHGTHGA RGTNGCYRTY

          ....|....| ....|....| ....|....| ....|....| ....|....|
             155        165        175        185        195
G kau     GTCGTCG**GCG GCGGCAACAT TTGGCG**TGGA AAAACGGGAA GCGAAATGGG
B thu     GTTGTTG**GTG GCGGTAACAT TTGGCG**TGGG AAAATTGGAA GTGAGATGGG
B cer     GTTGTTG**GTG GCGGTAACAT TTGGCG**TGGA AAAATTGGAA GTGAGATGGG
B hal     GTCGTAG**GTG GTGGCAACAT TTGGCG**CGGG ATGGCAGGTA GCGCGAAAGG
O ihe     GTTGTAG**GTG GAGGTAATAT TTGGAG**AGGT AAAGTCGGTA GTGAGATGGG
B sub     GTTGTAG**GCG GCGGCAACTT ATGGCG**CGGA AAAACAGGAA GTGACCTGGG
Conse     GTYGTHG**GYG GHGGYAAYWT WTGGMG**HGGD AWRRYNGGWA GYGMVMWRGG

          ....|....| ....|....| ....|....| ....|....| ....|....|
             205        215        225        235        245
G kau     CATGGACCGG GCGACGGCTG ATTACATGGG CATGTTGGCG ACGGTGATGA
B thu     CATGGATCGT GCAGGAGCAG ATTACATGGG CATGTTAGCG ACAGTTATGA
B cer     CATGGATCGT GCAGGAGCAG ATTACATGGG CATGTTAGCG ACAGTTATGA
B hal     GATGGACCGA GCAACTGCAG ATTACATGGG GATGCTAGCG ACGGTCATGA
O ihe     AATGGATCGT GCATCTGCTG ATTATATGGG GATGCTTGCT ACGATTATGA
B sub     CATGGACCGC GCGACTGCTG ACTATATGGG AATGCTGGCG ACAGTAATGA
Conse     VATGGAYCGN GCRDSDGCWG AYTAYATGGG VATGYTDGCK ACRRTNATGA
```

Figure 5.5. BLAST results for *smbA* gene (downsteam to *tsf*) used for reverse primer design (Primer designation site is indicated in bold characters)

G kau: *Geobacillus kaustophilus*, B thu: *Bacillus thuringiensis serovar konkukian*, B cer: *Bacillus cereus*, B hal: *Bacillus halodurans*, O ihe: *Oceanobacillus iheyensis*, B sub: *Bacillus subtilis*

```
          ....|....| ....|....| ....|....| ....|....| ....|....|
             255        265        275        285        295
G kau     ATGCCCTCGC CTTGCAGGAC AGCCTTGAGC AGCTCGGTGT GGAAACGCGG
B thu     ATTCATTAGC TCTTCAAGAT AGCTTGGAGA ATATCGGAAT TCAAACTCGC
B cer     ATTCATTAGC TCTTCAAGAT AGCTTGGAGA ATATCGGAAT TCAAACTCGC
B hal     ATTCATTAGC GTTACAGGAC AGCTTGGAAA ACCTTGACGT CCAATCACGC
O ihe     ACTCTCTTGC ATTACAAGAT GGCTTAGAAA CAATTGATGT TGAGACAAGA
B sub     ATTCGCTTGC TCTTCAAGAC AGCTTGGAAA CACTCGGAAT CCAGTCCAGA
Conse     AYKCNYTHGC NYTDCARGAY RGCYTDGARM MNMTYGRHRT BSARWCNMGV

          ....|....| ....|....| ....|....| ....|....| ....|....|
             305        315        325        335        345
G kau     GTGCAAACAT CGATTGAAAT GCGGCAAGTG GCCGAACCGT ACATTCGCCG
B thu     GTACAAACTT CAATTGAGAT GCGTCAAGTG GCAGAGCCTT ACATTCGTCG
B cer     GTACAAACTT CAATTGAGAT GCGTCAAGTG GCAGAGCCTT ACATTCGTCG
B hal     GTTCAAACCT CGATTGAAAT GAGACAAGTG GCAGAGCCTT ATATTCGCAG
O ihe     GTACAGACGT CCATTGAGAT GAGACAAGTA GCAGAACCTT ACATAAGAAG
B sub     GTGCAAACAT CCATTGAAAT GAGACAAGTT GCTGAACCGT ACATAAGAAG
Conse     GTDCARACNT CVATTGARAT GMGDCAAGTD GCHGARCCKT AYATWMGHMG

          ....|....| ....|....| ....|....| ....|....| ....|....|
             355        365        375        385        395
G kau     CCGGGCGATT CGCCATCTTG AGAAAAAACG AGTTGTCATT TTCGCCGCCG
B thu     TAAAGCAGTT CGTCACTTAG AGAAGAAACG TGTTGTTATC TTTGCAGCAG
B cer     TAAAGCAGTT CGTCACTTAG AGAAGAAACG TGTTGTTATC TTTGCAGCAG
B hal     ACGGGCGATT CGCCATTTGG AGAAAAAGCG GGTCGTCATT TTTGCCGCAG
O ihe     AAAAGCTATT CGCCATCTAG AGAAGAAACG TGTTGTCATT TTTGCAGCAG
B sub     AAAAGCGATA CGCCACTTAG AGAAAAAACG TGTCGTTATT TTCGCTGCGG
Conse     HMRRGCDRTW CGYCAYYTDG AGAARAARCG DGTYGTYATY TTYGCHGCVG

          ....|....| ....|....| ....|....| ....|....| ....|....|
             405        415        425        435        445
G kau     GCACAGGCAA CCCGTACTTT TCGACCGATA CGACCGCGGC GCTCAGGGCG
B thu     GTACAGGTAA TCCATACTTC TCTACAGATA CAACAGCAGC ATTACGTGCA
B cer     GTACAGGTAA TCCATACTTC TCTACAGATA CAACAGCAGC ATTACGTGCA
B hal     GAACAGGAAA TCCTTATTTC TCTACCGATA CAACGGCTGC ACTACGGGCA
O ihe     GAACGGGTAA TCCATATTTC TCTACAGACA CTACTGCCGC ATTAAGAGCA
B sub     GCACAGGAAA CCCATATTTC TCAACTGATA CGACAGCTGC ACTGCGGGCT
Conse     GHACRGGHAA YCCDTAYTTY TCDACHGAYA CDACNGCNGC RYTVMGDGCD

          ....|....| ....|....| ....|....| ....|....| ....|....|
             455        465        475        485        495
G kau     GCGGAAATCG AGGCGGACGT CATTTTAATG GCGAAAAACA ACGTCGATGG
B thu     GCAGAAATTG AAGCAGACGT AATTTTAATG GCGAAAAACA ATGTAGATGG
B cer     GCAGAAATTG AAGCAGACGT AATTTTAATG GCGAAAAACA ATGTAGATGG
B hal     GCTGAAATCG AGGCTGAAGT CATTTTAATG GCGAAAAATA AAGTGGACGG
O ihe     GCTGAAATTG AAGCTGAAGT AATTTTAATG GCAAAGAATA ATGTTGATGG
B sub     GCTGAAATTG AGGCAGACGT TATTTTAATG GCTAAAAATA ACGTTGACGG
Conse     GCDGAAATYG ARGCDGAMGT HATTTTAATG GCDAARAAYA AHGTNGAYGG
```

Figure 5.5. BLAST results for *smbA* gene (downsteam to *tsf*) used for reverse primer design (Primer designation site is indicated in bold characters)

G kau: *Geobacillus kaustophilus*, B thu: *Bacillus thuringiensis serovar konkukian*, B cer: *Bacillus cereus*, B hal: *Bacillus halodurans*, O ihe: *Oceanobacillus iheyensis*, B sub: *Bacillus subtilis* (continued)

```
          ....|....| ....|....| ....|....| ....|....| ....|....|
             505        515        525        535        545
G kau     CGTTTACAGC GCCGACCCGA ACGTCGACGC CAATGCGGTC AAATACGATG
B thu     CGTATATAAT GCGGATCCAT CTATTGACCC AACAGCTACG AAATACGAAA
B cer     CGTATATAAT GCGGATCCAT CTATTGACCC AACAGCTACG AAATACGAAA
B hal     TGTGTACAAT GCCGAACCTT CTGTTGATGT CAATGCCAAA AAATATACGT
O ihe     AGTATACACT GATGACCCTA AGTTAAACAA GTCAGCTAAG AAATACGAAT
B sub     TGTGTATAAT GCTGATCCTA GAAAAGATGA ATCAGCTGTT AAGTATGAAT
Conse     HGTDTAYAVY GMBGAHCCDW VNDWHRAYVH VWMWGCBRHN AARTAYRMDD

          ....|....| ....|....| ....|....| ....|....| ....|....|
             555        565        575        585        595
G kau     AGCTGTCGTA CTTGGACGTC ATCAAGCAAG GGCTCGGCGT CATGGATTCG
B thu     CGCTTACTTA CTTAGATGTA TTAAAAGAAG GTTTAGGTGT AATGGATTCT
B cer     CGCTTACTTA CTTAGATGTA TTAAAAGAAG GTTTAGGTGT AATGGATTCT
B hal     CCATTTCCTA TTTGGATGTG TTAAAAGAAG GCCTAGCGGT CATGGACTCA
O ihe     CGTTAACATA TCTTGAGATG TTAAATGAAG GATTAGGTGT AATGGATTCA
B sub     CACTTTCTTA TCTTGACGTT CTTAAAGACG GCCTGGAAGT CATGGATTCA
Conse     MVHTDWCNTA YYTDGABRTN HTHAADSAMG GNYTVGVNGT MATGGAYTCD

          ....|....| ....|....| ....|....| ....|....| ....|....|
             605        615        625        635        645
G kau     ACTGCTTCCT CGCTTTGCAT GGACAACAAC ATTCCGCTCA TCGTCTTCTC
B thu     ACAGCTTCTT CTCTATGTAT GGATAATGAT ATTCCATTAA TTGTATTCTC
B cer     ACAGCTTCTT CTCTATGTAT GGATAATGAT ATTCCATTAA TTGTATTCTC
B hal     ACGGCTTCAT CCCTATGTAT GGACAACAAC ATTCCGCTCA TTGTGTTCTC
O ihe     ACAGCATCTT CATTATGTAT GGATAATAAC ATTCCATTAA TTGTATTCTC
B sub     ACAGCTTCCT CTTTATGCAT GGACAATGAC ATTCCGCTTA TCGTCTTTTC
Conse     ACDGCWTCHT CNYTWTGYAT GGAYAAYRAY ATTCCRYTHA TYGTVTTYTC

          ....|....| ....|....| ....|....| ....|....| ....|....|
             655        665        675        685        695
G kau     GATTATGGAA GAAGGCAATA TTAAACGCGC TGTGTTAGGT GAAAACATCG
B thu     AGTTATGGAA AAAGGTAATA TTAAACGTGC CGTTTTAGGT GAAAATATCG
B cer     AGTTATGGAA AAAGGTAATA TTAAACGTGC CGTTTTAGGT GAAAATATCG
B hal     GATTATGGAA GAAGGTAATA TTAAAAAGGC CGTTCTTGGA GAAGAAATCG
O ihe     AATTATGGAA GAAGGAAACA TTAAACGGGT CGTTCAAGGT GAGACGATCG
B sub     TATTATGGAA GAAGGAAATA TCAAACGTGC CGTTATCGGT GAATCAATCG
Conse     DRTTATGGAA RAAGGHAAYA TYAAAMRBGY YGTKHWHGGW GARDMNATCG

          ....|....| ....|....| ....|..
             705        715        725
G kau     GAACGATCGT AAGGGGGAAA TAA----
B thu     GAACAGTTGT AAGGGGGAAA TAA----
B cer     GAACAGTTGT AAGGGGGAAA TAA----
B hal     GTACTGTTGT AAGGGGGAG- TAACTAA
O ihe     GAACTACAAT AAGGGGGAAA TAA----
B sub     GAACGATCGT GAGGGGGAAA TAA----
Conse     GWACDRYHRT RAGGGGGARA TAACTAA
```

Figure 5.5. BLAST results for *smbA* gene (downsteam to *tsf*) used for reverse primer design (Primer designation site is indicated in bold characters)

G kau: *Geobacillus kaustophilus*, B thu: *Bacillus thuringiensis serovar konkukian*, B cer: *Bacillus cereus*, B hal: *Bacillus halodurans*, O ihe: *Oceanobacillus iheyensis*, B sub: *Bacillus subtilis* (continued)

5.2.2. PCR Amplifications of the Complete *Geobacillus anatolicus tsf* Gene

By using degenerate primer pair rpsBF and smbAR *Geobacillus anatolicus*, a genomic DNA fragment carrying the *tsf* gene was amplified (Figure 5.6). The expected product size of the PCR fragment ~1200 bp was verified. For an additional confirmation, the PCR product amplified by using rpsBF and smbAR primers was purified and this fragment was used as DNA template for PCR with degenerate primer pair TsF45 and TsR696 (targetting the *tsf* gene), to confirm that the 1200 bp fragment indeed contains the *tsf* gene (Figure 5.7).

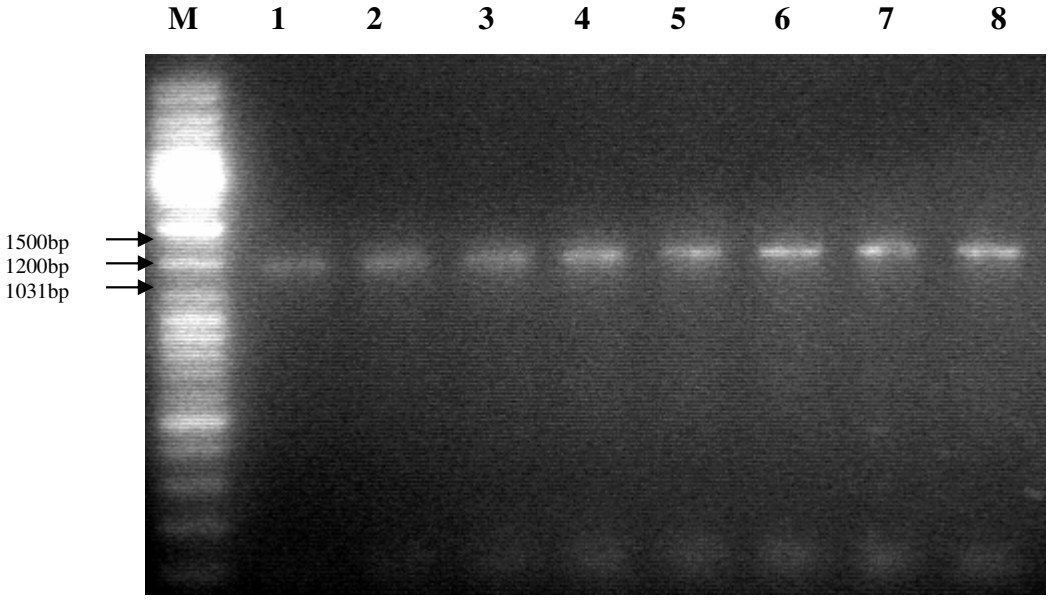

Figure 5.6. PCR amplification of *Geobacillus anatolicus* DNA fragment carrying the *tsf* gene by using primer pair rpsBF and smbAR.
M, molecular size marker; 1-8, annealing temperature gradient (50, 50.7, 52, 53.7, 56.2, 58.1, 59.3, 60^{o}C respectively).

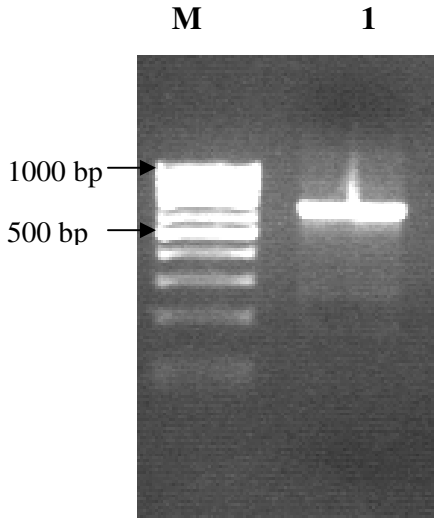

Figure 5.7. Control PCR for verifying the rpsBF /smbAR amplified fragment contains the *tsf* gene

M, molecular size marker; 1, PCR product amplifed by TsF45 and TsR696 primer pair

5.2.3. *Geobacillus anatolicus tsf* Gene Sequence

The primary sequence information obtained from the sequencing of the PCR product amplified by TsF45/TsR696 was used to design the two non-degenerate primers TsF592 and TsR357 (Table 3.3). The PCR product amplifed by using rpsBF /smbAR primers carrying the *Geobacillus anatolicus tsf* gene were sequence by using the primers TsF45, TsR696, TsF592, TsR357, rpsBF and smbAR . The complete *tsf* gene sequence obtained this way is given in Figure 5.8. *Geobacillus anatolicus tsf* gene contains 885 bases including the translational start and the stop site of the protein. Nucleic Acid Statistics (NASTATS) were given in Table 5.4.

Shine-Dalgarno sequence start codon

```
   1 AAGAGAAGGA GGATCTTTTT ATGGCGATTA CCGCACAAAT GGTAAAAGAG CTGCGCGAAA
  61 AAACGGGCGC AGGCATGATG GATTGCCAAA AGGCGCTCAC CGAAACGAAC GGTGACATGG
 121 AAAAAGCGAT CGACTGGCTG CGCGAAAAAG GAATTGCCAA AGCGGCAAAA AAGCGGACCG
 181 CATCCCCGGC GGAAGGAATG ACATACATCG CTGTGGAGGC CACTGCGGCC GTCATTTTGG
 241 AAGTGAATTC GGAAACGGAC TTCGTTGCCA AAAACGAAGC GTTCCAAACG CTCGTTAAGG
 301 AGCTGGCTGC GCATTTGCTG AAACAAAAAC CGGCTTCGCT TGATGAAGCG CTCGGACAAA
 361 CGATGGACAA CGGTTCCACT GTCCAAGATT ATATTAACGA AGCGATCGCG AAAATCGGCG
 421 AAAAAATCAC GCTCCGCCGC TTTGCTGTCG TCAACAAAGC GGACGGCGAA ACGTTTGGCG
 481 CGTACTTGCA CATGGGCGGG CGCATCGGCG TATTAACATT ATTAGCCGGC AATGCAAGCG
 541 AAGAGATCGC CAAAGATGTG GCCATGCATA TCGCTGCGCT CCATCCGAAA TATGTTTCAC
 601 GCGATGAAGT GCCGCAAGAA GAGATTGCGC GCGAACGCGA AGTGTTGAAA CAACAAGCGT
 661 TAAATGAAGG CAAGCCGGAA AAAATCGTTG AGAAAATGGT CGAAGGCCGG CTGAACAAGT
 721 TTTACGAAGA CGTTTGCCTG CTTGAGCAAG CGTTCGTGAA AAACCCGGAT GTGACGGTGC
 781 GCCAATACGT CGAGTCGAAC GGAGCAACTG TGAAGCAGTT CATCCGCTAC GAAGTCGGCG
 841 AAGGGCTCGA AAAACGTCAA GATAATTTCG CCGAAGAAGT CATGAGCCAA GTAAGAAAGC
 901 AATGA
```

stop codon

Figure 5.8. Nucleic acid sequence of *Geobacillus anatolicus tsf* gene

Table 5.4. NASTATS search result of *Geobacillus anatolicus tsf* gene

	Total	**Percentage**
Adenine (A)	275	% 31.1
Thymine (T)	156	% 17.6
Cytosine (C)	203	% 22.9
Guanine (G)	251	% 28.4
A and T	431	% 48.7
C and G	454	% 51.3

The nucleotide sequence of the *Geobacillus anatolicus tsf* gene was used to deduce the aminoacid sequence of the EF-Ts protein. The complete aminoacid sequence of the *Geobacillus anatolicus* EF-Ts, obtained by using the SIX-FRAME software is given in Figure 5.9 and Figure 5.10.

```
      M  A  I  T  A  Q  M  V  K  E  L  R  E  K  T  G  A  G  M  M
1     atggcgattaccgcacaaatggtaaaagagctgcgcgaaaaaacgggcgcaggcatgatg  60
      D  C  Q  K  A  L  T  E  T  N  G  D  M  E  K  A  I  D  W  L
61    gattgccaaaaggcgctcaccgaaacgaacggtgacatggaaaaagcgatcgactggctg  120
      R  E  K  G  I  A  K  A  A  K  K  R  T  A  S  P  A  E  G  M
121   cgcgaaaaaggaattgccaaagcggcaaaaaagcggaccgcatccccggcggaaggaatg  180
      T  Y  I  A  V  E  A  T  A  A  V  I  L  E  V  N  S  E  T  D
181   acatacatcgctgtggaggccactgcggccgtcattttggaagtgaattcggaaacggac  240
      F  V  A  K  N  E  A  F  Q  T  L  V  K  E  L  A  A  H  L  L
241   ttcgttgccaaaaacgaagcgttccaaacgctcgttaaggagctggctgcgcatttgctg  300
      K  Q  K  P  A  S  L  D  E  A  L  G  Q  T  M  D  N  G  S  T
301   aaacaaaaaccggcttcgcttgatgaagcgctcggacaaacgatggacaacggttccact  360
      V  Q  D  Y  I  N  E  A  I  A  K  I  G  E  K  I  T  L  R  R
361   gtccaagattatattaacgaagcgatcgcgaaaatcggcgaaaaaatcacgctccgccgc  420
      F  A  V  V  N  K  A  D  G  E  T  F  G  A  Y  L  H  M  G  G
421   tttgctgtcgtcaacaaagcggacggcgaaacgtttggcgcgtacttgcacatgggcggg  480
      R  I  G  V  L  T  L  L  A  G  N  A  S  E  E  I  A  K  D  V
481   cgcatcggcgtattaacattattagccggcaatgcaagcgaagagatcgccaaagatgtg  540
      A  M  H  I  A  A  L  H  P  K  Y  V  S  R  D  E  V  P  Q  E
541   gccatgcatatcgctgcgctccatccgaaatatgtttcacgcgatgaagtgccgcaagaa  600
      E  I  A  R  E  R  E  V  L  K  Q  Q  A  L  N  E  G  K  P  E
601   gagattgcgcgcgaacgcgaagtgttgaaacaacaagcgttaaatgaaggcaagccggaa  660
      K  I  V  E  K  M  V  E  G  R  L  N  K  F  Y  E  D  V  C  L
661   aaaatcgttgagaaaatggtcgaaggccggctgaacaagttttacgaagacgtttgcctg  720
      L  E  Q  A  F  V  K  N  P  D  V  T  V  R  Q  Y  V  E  S  N
721   cttgagcaagcgttcgtgaaaaacccggatgtgacggtgcgccaatacgtcgagtcgaac  780
      G  A  T  V  K  Q  F  I  R  Y  E  V  G  E  G  L  E  K  R  Q
781   ggagcaactgtgaagcagttcatccgctacgaagtcggcgaagggctcgaaaaacgtcaa  840
       D  N  F  A  E  E  V  M  S  Q  V  R  K  Q  *
841   gataatttcgccgaagaagtcatgagccaagtaagaaagcaatga  885
```

Figure 5.9. The SIX-FRAME result of *Geobacillus anatolicus tsf* gene

```
  1 MAITAQMVKE LREKTGAGMM DCQKALTETN GDMEKAIDWL REKGIAKAAK KRTASPAEGM
 61 TYIAVEATAA VILEVNSETD FVAKNEAFQT LVKELAAHLL KQKPASLDEA LGQTMDNGST
121 VQDYINEAIA KIGEKITLRR FAVVNKADGE TFGAYLHMGG RIGVLTLLAG NASEEIAKDV
181 AMHIAALHPK YVSRDEVPQE EIAREREVLK QQALNEGKPE KIVEKMVEGR LNKFYEDVCL
241 LEQAFVKNPD VTVRQYVESN GATVKQFIRY EVGEGLEKRQ DNFAEEVMSQ VRKQ
```

Figure 5.10. The amino acid sequence of *Geobacillus anatolicus* EF-Ts

The amino acid sequence of *Geobacillus anatolicus tsf* was compared with other known thermophilic *tsf* sequences in the Bacterial GenBank using BLASTN search at Biology WorkBench. Conserved features are given in Figure 5.11.

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                    5         15         25         35         45
Tmar            ------MEIS MDLIKKLREM TGAGILDCKK ALEEANGDME KAVEILRKKG
Aaeo            ----MRMAVS MEDVKKLREM TGAGMLDCKK ALEEAGGDIE KAKEILRVKG
Tthe            -------MSQ MELIKKLREA TGAGMMDVKR ALEDAGWDEE KAVQLLRERG
Telo            -----MAEIS AKLVKELRDK TGAGMMDCKK ALQESNGDME AAITWLRQKG
Tten            -------MIS AQAVKELRER TGAGMMDCKN ALIEANGDIE KAIDILREKG
Sthe1           -----MAEIT AKMVAELRAR TGAGMMDCKK ALMETGGDFD KAVDWLREKG
Sthe2           MEKNKMAEIT AKLVKELREK SGAGVMDAKK ALVEVDGDIE KAIELLREKG
Tfus            -----MANYT AADVKRLREL TGAGMMACKK ALEESGGDFD KAIEALRIKG
Gkau            ------MAIT AQMVKELREK TGAGMMDCKK ALTETNGDME KAIDWLREKG
Gana            ------MAIT AQMVKELREK TGAGMMDCQK ALTETNGDME KAIDWLREKG
Consensus                     LR    GAG       AL     D    A   LR  G

                ....|....| ....|....| ....|....| ....|....| ....|....|
                    55         65         75         85         95
Tmar            AATAEKKAGR TTKEGIIVAY VHFNGRIGVL LEMNCETDFV ARTDEFKELA
Aaeo            LAKAEKKAGR ETKEGLIYVI VSEDRKKGAM IELNCETDFV ARNEEFRKLA
Tthe            AMKAAKKADR EAREGIIGHY IHHNQRVGVL VELNCETDFV ARNELFQNLA
Telo            LASAGKKAGR VTSEGLVDSY IHTGGRIGVL VEVNCETDFV ARNEKFKTLV
Tten            LAAAAKKAGR TANEGLVEAY IHGGGRIGVL VEVNCETDFV ANTEEFRNFV
Sthe1           LAAAAKKAGR VAAEGRVHAI VEDGARHGVL VEVNCETDFV ARGEAFINLC
Sthe2           MAKAAKKADR IAAEGLTGIY VSG--NVAAV VEVNAETDFV AKNAQFVELV
Tfus            AKDVGKRAER TAANGLIALA QDGD-TSAVL LELNCETDFV AKNDKFQELA
Gkau            IAKAAKKADR IAAEGMAYIA VEG--NTAVI LEVNSETDFV AKNEAFQTLV
Gana            IAKAAKKRTA SPAEGMTYIA VEA--TAAVI LEVNSETDFV AKNEAFQTLV
Consensus         K         G                    E N ETDFV A    F

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   105        115        125        135        145
Tmar            Y--------- ---------- ---------- ---------- ----------
Aaeo            ERITRHILEK DENKNKSGEG SEILSQELYD EPGKTVETLI KEAIAKIGEN
Tthe            K--------- ---------- ---------- ---------- ----------
Telo            Q--------- ---------- ---------- ---------- ----------
Tten            K--------- ---------- ---------- ---------- ----------
Sthe1           DHVARVILQA RPASLEALQE ALG------- -DTVKEAVAK IGENIQVRRF
Sthe2           NETAKVIAEG KPANNEEALA LTMP------ -SGETLEAAY VTATATIGEK
Tfus            AELAGFVART SPSDVPSLLS ADYA------ -DGKTVSQVI EELSAVIGEK
Gkau            KELAAHLLKQ KPASLDEALG QTMD------ -NGSTVQDYI NEAIAKIGEK
Gana            KELAAHLLKQ KPASLDEALG QTMD------ -NGSTVQDYI NEAIAKIGEK
Consensus
```

Figure 5.11. Aminoacid sequence alignment of some thermophilic elongation factors Tmar: *Thermotoga maritima*, Aaeo: *Aquifex aeolicus*, Tthe: *Thermus thermophilus*, Telo: *Thermosynechococcus elongatus*, Tten: *Thermoanaerobacter tengcongensis*, Sthe1: *Sybiobacterium thermophilum*, Sthe2: *Streptococcus thermophilus*, Tfus: *Thermobifida fusca*, Gkau: *Geobacillus kaustophilus*, Gana: *Geobacillus anatolicus*

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                   155        165        175        185        195
Tmar            ---------- ---------- ---------- ---------- -----NLAKQ
Aaeo            IRLSRYCRYD TEDYLHSYVH GGGRIGVLLD FKAPELNDQV LRLVQDVAMQ
Tthe            ---------- ---------- ---------- ---------- -----DLAMH
Telo            ---------- ---------- ---------- ---------- -----DIAKQ
Tten            ---------- ---------- ---------- ---------- -----EICMQ
Sthe1           ERYEVADAGR VHAYIHGDGR VGVLIELTTA TPEVAAHPEV EALCHELALQ
Sthe2           ISLRRFAVVE KTDAQHFG-- AYQHNGGRIG VVSVIEG-GD EAIAKQISMH
Tfus            IELRRFAKVE GAYVASYMHK SDPDLPPTLG VLVELDK-PN AEVAKDLAQQ
Gkau            ITLRRFAVVN KADGETFG-- AYLHMGGRIG VLTLLAGNAS EDVAKDVAMH
Gana            ITLRRFAVVN KADGETFG-- AYLHMGGRIG VLTLLAGNAS EEIAKDVAMH
Consensus

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   205        215        225        235        245
Tmar            VAAM-K---- ---------- ---------- ---------- -------PLY
Aaeo            IAAM-R---- ---------- ---------- ---------- -------PEY
Tthe            IAMM-N---- ---------- ---------- ---------- -------PRY
Telo            IAACPN---- ---------- ---------- ---------- -------VEF
Tten            IAAA-N---- ---------- ---------- ---------- -------PKY
Sthe1           IASM-R---- ---------- ---------- ---------- -------AQY
Sthe2           IAAMKPTVLS YSELDEQFVK DELAQLNHAI DQDNESRAMV NKPALPHLKY
Tfus            IAALAP---- ---------- ---------- ---------- --------KY
Gkau            IAALHP---- ---------- ---------- ---------- --------KY
Gana            IAALHP---- ---------- ---------- ---------- --------KY
Consensus        A

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   255        265        275        285        295
Tmar            VRREDVPAEV IEKEKEIYR- -AQIKDKPEN IVEKIVEGKL EKFFEQ----
Aaeo            VRIEDIPQEV LERERRILRE QALQEGKPEH IVDKIVEGKL KKFYQE----
Tthe            VSAEEIPAEE LEKERQIYIQ AALNEGKPQQ IAEKIAEGRL KKYLEE----
Telo            VSLDDIPAEY KEKERQIALG SDALKGKPPE VQEKIVAGKL EKTLKE----
Tten            ISKEDVPQEV LEKEKEILRA QALNEGKPAN VIDRIVEGRL EKFYKE----
Sthe1           VRPEDVPAQV IEHEREILKA QAINEGKKPE IAEKMVAGRI QKFFQE----
Sthe2           GSKAQLTDEV IAQAEEDIKA ELAAEGKPEK IWDKIIPGKM DRFILDNTKV
Tfus            ISRDDVPADV VENERRIAKA TAREEGKPEQ ALPKIVEGRL NGFFKD----
Gkau            VSRDDVPQEE IAHEREVLKQ QALNEGKPEK IVEKMVEGRL NKFYED----
Gana            VSRDEVPQEE IARE REVLKQ QALNEGKPEK IVEKMVEGRL NKFYED----
Consensus                                    K          G
```

Figure 5.11. Aminoacid sequence alignment of some thermophilic elongation factors Tmar: *Thermotoga maritima*, Aaeo: *Aquifex aeolicus*, Tthe: *Thermus thermophilus*, Telo: *Thermosynechococcus elongatus*, Tten: *Thermoanaerobacter tengcongensis*, Sthe1: *Sybiobacterium thermophilum*, Sthe2: *Streptococcus thermophilus*, Tfus: *Thermobifida fusca*, Gkau: *Geobacillus kaustophilus*, Gana: *Geobacillus anatolicus* (continued)

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                   305        315        325        335        345
Tmar         ---ACLYEQT YIFDDTKKVK DLINELIAKT GENIRVSRFT RYEIGEGYED
Aaeo         ---KVLLEQP FIKEEKKQVK DVIKES---- GLNVEIKRFC RFELGGL---
Tthe         ---VVLLEQP FVKDDKVKVK ELIQQAIAKI GENIVVRRFC RFELGA----
Telo         ---LCLLYQP FIRDQSKTVE ELVKEHIAEL GENIRIRRFQ RFVLGEGIEK
Tten         ---NCLLEQE YIRDPEKTVK DLLNEMIAKL GENIVIRRFA RFERGEGIEK
Sthe1        ---VCLLEQE WVKDSKKTVG ALVKEVAGKV GGEITVKRFV RYEKGEGIEK
Sthe2        DQAYTLLAQV YIMDDSKTVE AYLESV---- --NASVVEFV RFEVGEGIEK
Tfus         ---ATLLGQP FVKDNKKTIQ QVVDEA---- --GVTVRRFV RFKVGQA---
Gkau         ---VCLLEQA FVKNPDVTVR QYVESN---- --GATVKQFI RYEVGEGLEK
Gana         ---VCLLEQA FVKNPDVTVR QYVESN---- --GATVKQFI RYEVGEGLEK
Consensus         L  Q                                F  R   G

                ....|....| ....|....| ....|....| ....|....| ....|....|
                   355        365        375        385        395
Tmar         ---------- ---------- ---------- ---------- ----------
Aaeo         ---------- ---------- ---------- ---------- ----------
Tthe         ---------- ---------- ---------- ---------- ----------
Telo         QETNLAEEVA AQTQAMRAAA QTAAAAETAP PEVSEPEPAA AVTAEEPTPE
Tten         ---------- ---------- ---------- ---------- ----------
Sthe1        RQDDFAEEVM RQAGLK---- ---------- ---------- ----------
Sthe2        ASNDFESEVA ATMAAALNN- ---------- ---------- ----------
Tfus         ---------- ---------- ---------- ---------- ----------
Gkau         RQDNFAEEVM SQVRKQ---- ---------- ---------- ----------
Gana         RQDNFAEEVM SQVRKQ---- ---------- ---------- ----------
Consensus

                ....|....| ....|....| ....|....| ....|....| ....
                   405        415        425        435
Tmar         ---------- ---------- ---------- ---------- ----
Aaeo         ---------- ---------- ---------- ---------- ----
Tthe         ---------- ---------- ---------- ---------- ----
Telo         PVAAAEQPAE PVSELVEQPG AESKGFGAAT KKSGGKSRSN KKKK
Tten         ---------- ---------- ---------- ---------- ----
Sthe1        ---------- ---------- ---------- ---------- ----
Sthe2        ---------- ---------- ---------- ---------- ----
Tfus         ---------- ---------- ---------- ---------- ----
Gkau         ---------- ---------- ---------- ---------- ----
Gana         ---------- ---------- ---------- ---------- ----
Consensus
```

Figure 5.11. Aminoacid sequence alignment of some thermophilic elongation factors Tmar: *Thermotoga maritima*, Aaeo: *Aquifex aeolicus*, Tthe: *Thermus thermophilus*, Telo: *Thermosynechococcus elongatus*, Tten: *Thermoanaerobacter tengcongensis*, Sthe1: *Symbiobacterium thermophilum*, Sthe2: *Streptococcus thermophilus*, Tfus: *Thermobifida fusca*, Gkau: *Geobacillus kaustophilus*, Gana: *Geobacillus anatolicus* (continued)

5.3. Cloning of the *Geobacillus anatolicus tsf* Gene

5.3.1. Cloning Strategy

pCR®T7/NT-TOPO® vector (see Figure 3.2) was chosen for the cloning of the *Geobacillus anatolicus tsf* gene. The pCR®T7/NT-TOPO® plasmid is used for cloning of Taq-amplified PCR products directly into a prokaryotic expression vector. The plasmid carries the bacterial phage T7 promoter for high-level protein expression in *Escherichia coli*. pCR®T7/NT-TOPO® contains; N-terminal Xpress™ epitope for detection with an Anti-Xpress™ Antibody, N-terminal polyhistidine (6xHis) tag for purification using ProBond™ resin and detection with an Anti-HisG Antibody and enterokinase cleavage site for efficient removal of the N-terminal fusion tag. The vector also contains an ampicilin resistance site as a selection marker after transformation. However, the original pCR®T7/NT-TOPO® was not used in the cloning experiments. Instead, the vector was shortened between Nde I cutting site and Csp45 I (BstB I) cutting site (see Figure 3.2 and Figure 5.12). By this way, the non-coding fragment which includes the Xpress epitope, the enterokinase cleavage site and the His-tag together with the first ATG codon was removed. This modified vector was called pT7D3.

Figure 5.12. Cloning site of the plasmid pCR®T7/NT-TOPO® before digestions

5.3.2. Constructing Primers for Cloning of *tsf* Gene

In order to insert the *tsf* gene into the pT7D3 vector (the shortened form of the pCR®T7/NT-TOPO®) a short sequence needed to be added to the downstream and to the upstream of the PCR product carrying the *tsf* gene. For this purpose, two primers were designed. A forward primer, TsF2, contained a Nde I cleavage site as an overhanging region of the primer. This primer also includes the beginning of the gene including the start codon inside the cleavage site. A reverse primer, TsR2, contained the very end sequence of

the gene including the stop codon and a Csp45 I (BstB I) cutting site as the overhanging region of the primer (Figure 5.13, see also Materials and Methods).

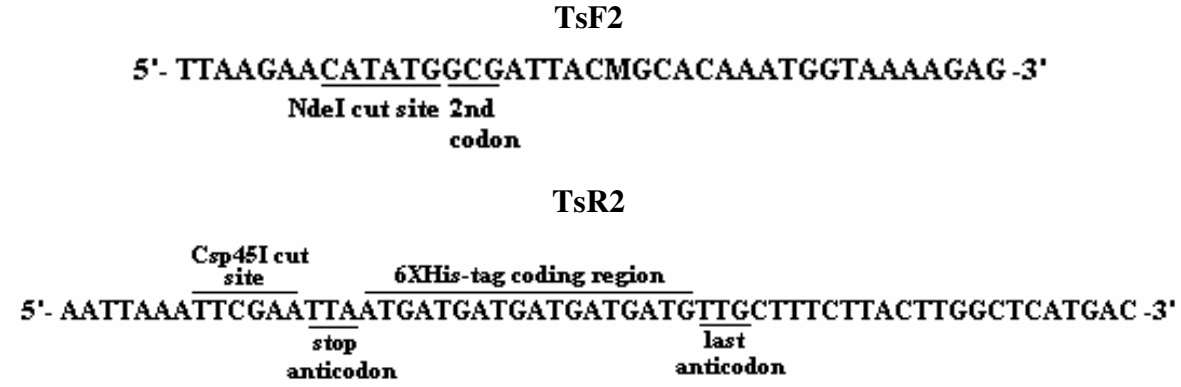

Figure 5.13. Primers constructed for PCR cloning according to the *tsf* full sequence data and the plasmid restriction sites

PCR amplification using TsF2 and TsR2 primers with *Geobacillus anatolicus* genomic DNA gave a unique PCR product at every annealing temperature between 50.0°C to 60.0°C as analyzed by electrophoresis on 1 % agarose. Figure 5.14 shows the result of this amplification at 57°C after the PCR product is purified for restriction enzyme cleavage.

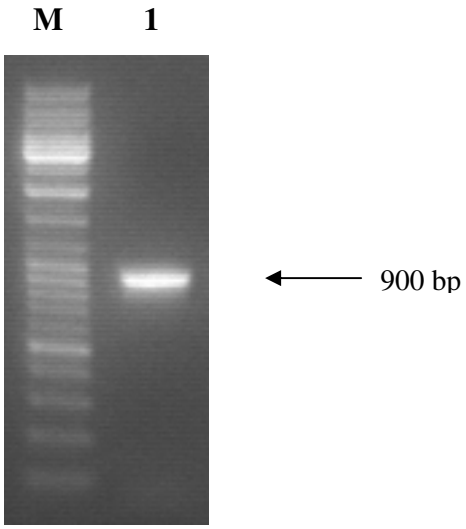

Figure 5.14. PCR amplification of *Geobacillus anatolicus tsf* gene from genomic DNA using primer pair TsF2 and TsR2
M, molecular size marker; 1, PCR product

5.3.3. Construction of pT7D3Ts

pCR®T7/NT-TOPO® vector was digested with Nde I and Csp45 I. The resulting linear vector was named as pT7D3. pT7D3 vector was ligated with the Nde I and Csp45 I digested PCR fragment which was amplified by TsF2 and TsR2 primers as described above (See Figure 5.16, Figure 5.17, Figure 5.18). The resulting circular plasmid was named as pT7D3Ts. pT7D3Ts was used to transform *Escherichia coli* JM109 (DE3) strain. Plasmid purification from ampicilin resistant transformants was described in section 4.6. Purified plasmids were analyzed on 1% agarose gel as shown in Figure 5.15.

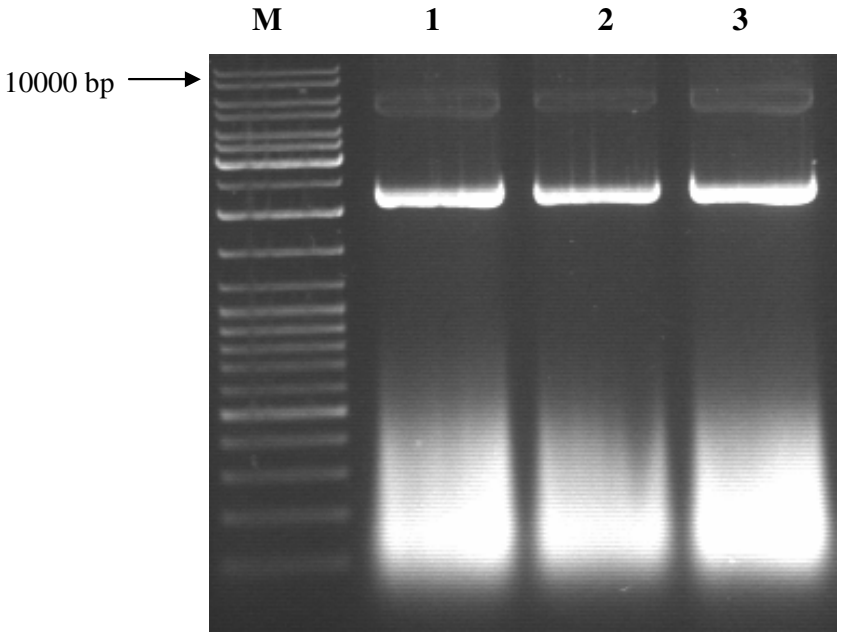

Figure 5.15. Results for plasmid purification from ampicilin resistant colonies
M, molecular size marker; 1-3, plasmids bearing the *tsf* gene

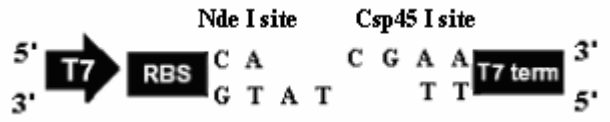

Figure 5.16. Cloning site created on pCR®T7/NT-TOPO® after restriction digestions during the construction of pT7D3

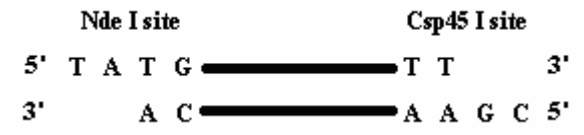

Figure 5.17. PCR product containing *tsf* gene after restriction digestions

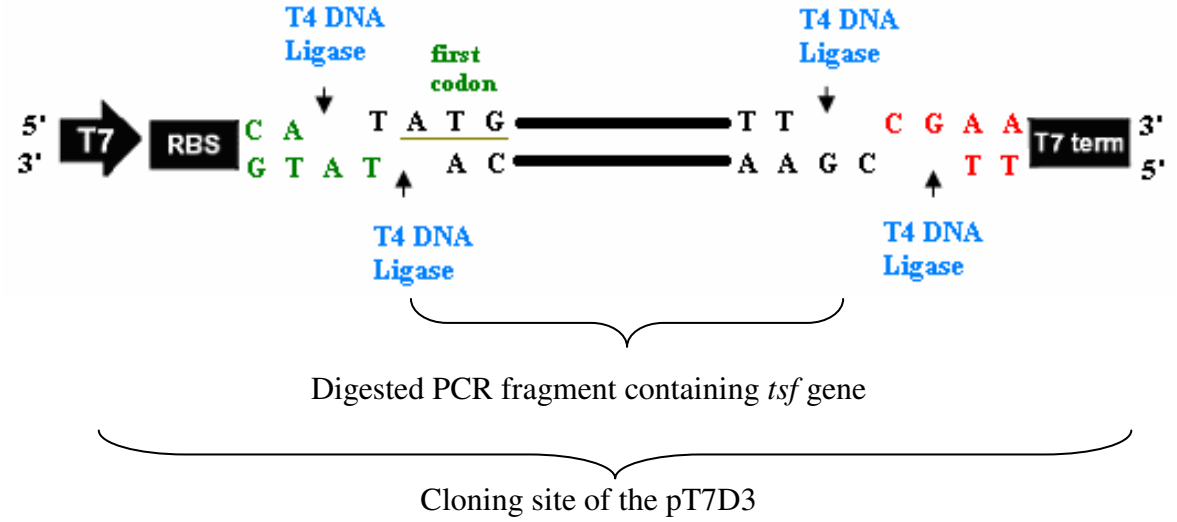

Figure 5.18. Cloning strategy (formation of pT7D3Ts from pT7D3)

5.3.4. Verification of *tsf* insertion into pT7D3

In order to confirm that the PCR fragment containing the *tsf* gene is actually inserted into the plasmid pT7D3, the following control experiments were made. First, the plasmid was digested with Nde I and Csp45 I and the expected fragment sizes confirmed the insertion of the *tsf* gene (Figure 5.19). Second, the purified plasmid was used to amplify a DNA fragment by using the primer pair TsF2 and TsR2. The product size of this fragment also confirmed the correct insertion of the *tsf* gene. The result of this experiment is shown in Figure 5.20. The plasmid, pT7D3Ts, was also sequenced with the plasmids own primers, T7 Forward and pRSET Reverse, which were available together with the plasmid. The sequence results also confirmed the insert was the *tsf* gene.

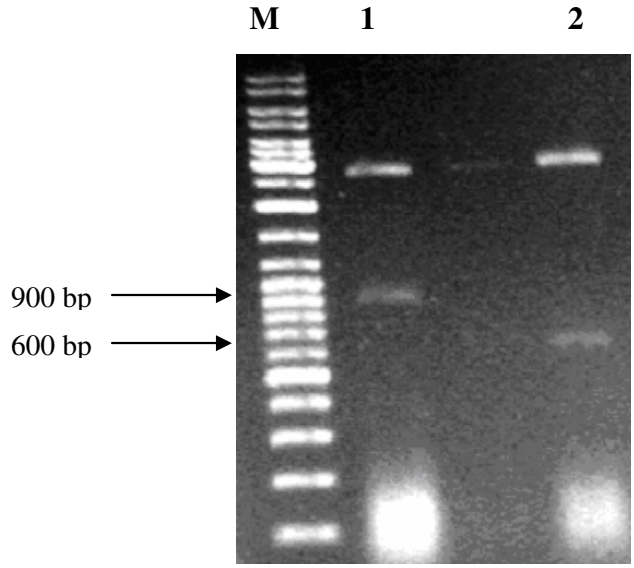

Figure 5.19. Restriction enzyme digestions of the purified plasmids
M, Molecular size marker; 1, Nde I and Csp45 I double digestion of the recombinant plasmid pT7D3Ts; 2, Nde I and Csp45 I double digestion of the original plasmid pCR®T7/NT-TOPO®

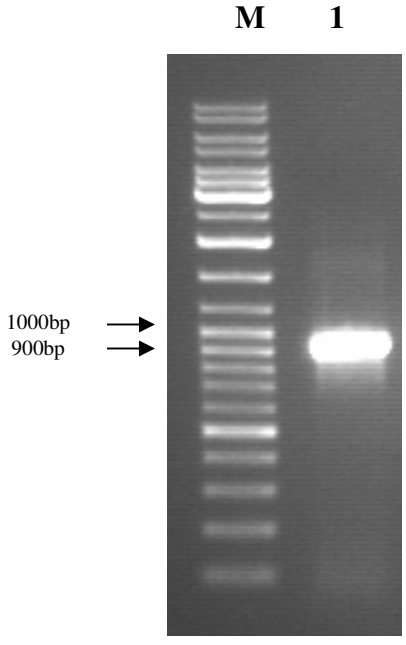

Figure 5.20. Amplification of the *tsf* gene from purified plasmid pT7D3Ts using primer pair TsF2 and TsR2
M, Molecular size marker; 1, resulting PCR product

5.4. Expression

Escherichia coli JM109(DE3) strain transformed with plasmid pT7D3Ts carrying the *tsf* gene was used for protein expression. This strain contains T7 RNA polymerase gene which is required for the expression of the recombinant proteins since the recombinant gene is under the control of the T7 promoter. The expression of the *tsf* gene was induced by IPTG. As a control, cells containing the plasmid pCR®T7/NT-TOPO® (lacking *tsf*) were used. The expression of *tsf* gene was analyzed on 10 % SDS-PAGE gel as a protein band at 34 kDa, which was approximately the same as the calculated molecular weight of the protein. The amount of protein increased after IPTG induction for the cells carrying the pT7D3Ts plasmid. A band of approximately the same size was observed in the control cells carrying pCR®T7/NT-TOPO® (lacking *tsf*) but its intensity stayed constant after IPTG induction (Figure 5.21). For preparative expression, cells were harvested 4 hours after IPTG induction.

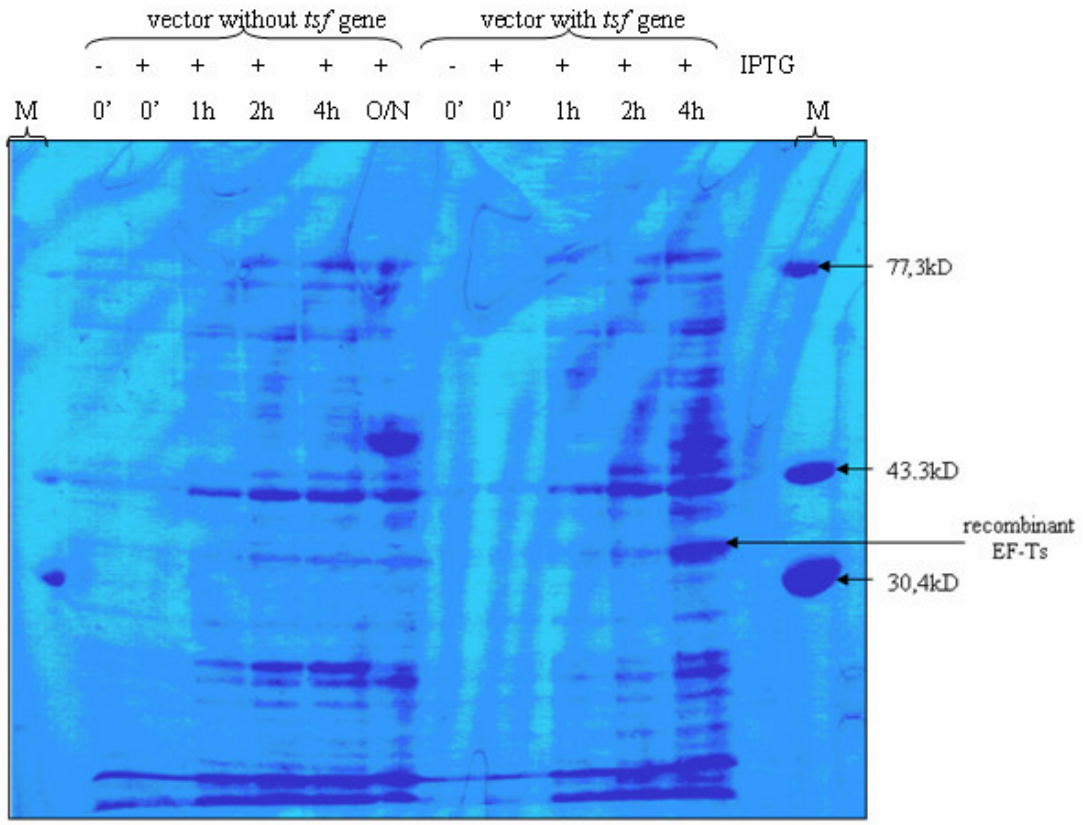

Figure 5.21. SDS results for recombinant plasmid expression
M, Protein standard containing *Escherichia coli* EF-Tu, EF-Ts, EF-G

5.5. Ni-Affinity Column Chromatography

Ni-affinity chromotography enables purification of 6x His-tag containing proteins. *Escherichia coli* JM109(DE3) cells harboring pT7D3Ts were grown in LB medium in the presence of 150 mg/ml ampicillin and the *tsf* gene expression was induced by IPTG. After 4 h induction, cells were harvested and the cell lysate was applied directly onto the Ni-affinity column. The recombinant His-tagged protein was eluted by an imidazole gradient between 25 mM to 150 mM. *Geobacillus anatolicus* EF-Ts protein was eluted at approximately 100 mM imidazole. Elution and the purity of the EF-Ts protein was monitored by SDS-PAGE (See Figure 5.22). Purified recombinant *Geobacillus anatolicus* EF-Ts is shown in Figure 5.23. Total protein yield as determined by Bradford assay (section 4.11) was 1.95 mg from 1.6 liter of *Escherichia coli* culture.

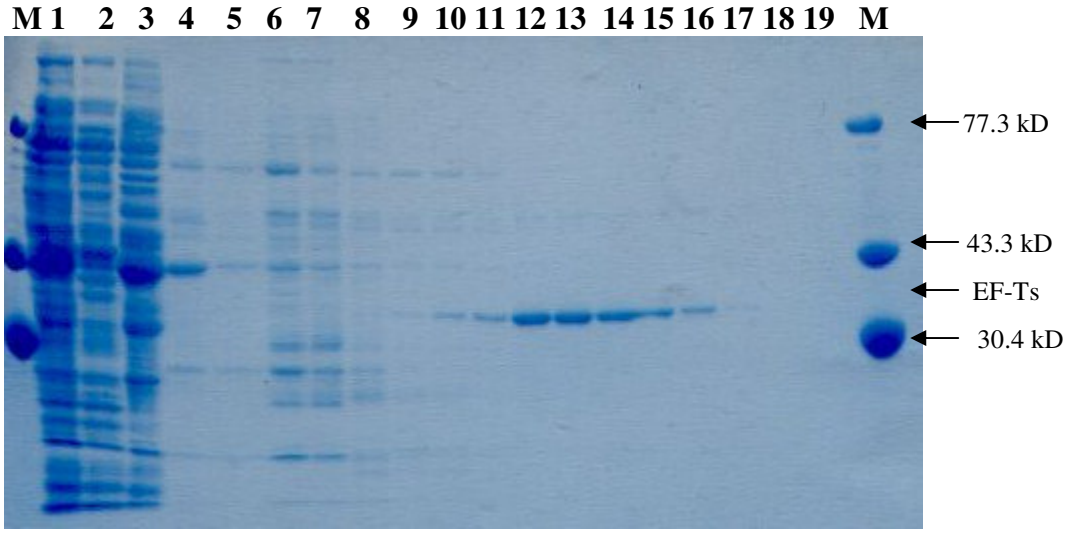

Figure 5.22. SDS results for Ni-Affinity chromotography of overexpression of *Geobacillus anatolicus* EF-Ts gene

M, Protein standard containing *Escherichia coli* EF-Tu, EF-Ts, EF-G; 1, sample before loading to column, 2-19, varying numbers of elution according to imidazole gradient

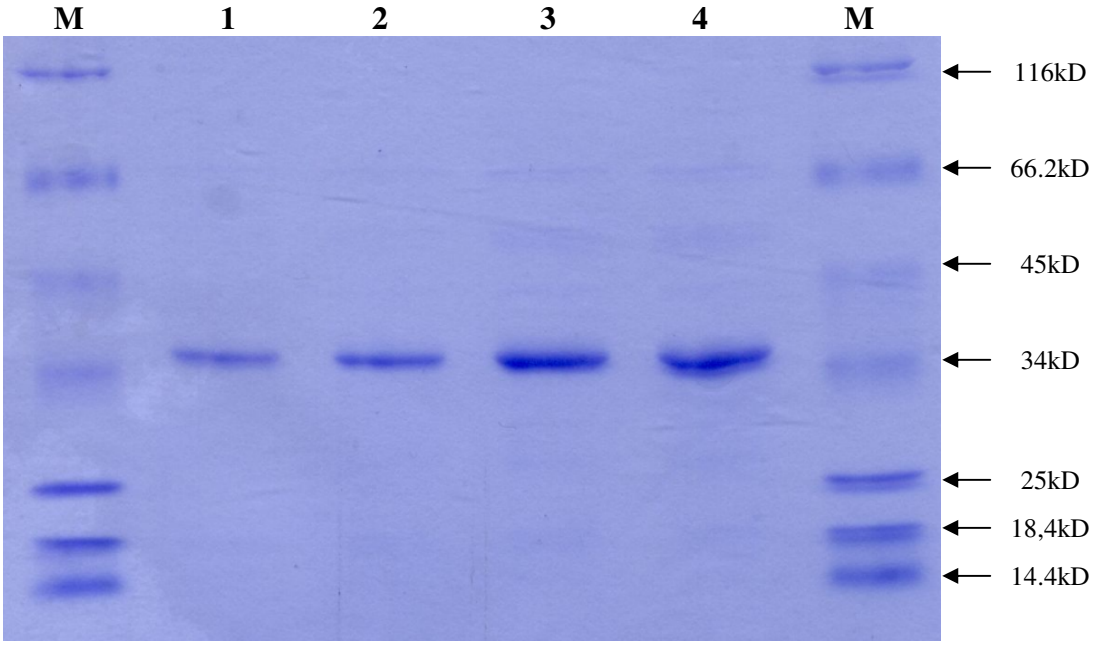

Figure 5.23. Purified recombinant *Geobacillus anatolicus* EF-Ts
M, Protein standard; 1-4, recombinant *Geobacillus anatolicus* EF-Ts loaded 1µl, 2µl, 3µl, 4µl respectively

5.6. Binding assay for *Escherichia coli* EF-Tu and *Geobacillus anatolicus* EF-Ts

Recombinant *Geobacillus anatolicus* EF-Ts was found to be fully active in *Escherichia coli* EF-Tu binding as analyzed by mobility shift assay (see section 4.12). *Geobacillus anatolicus* EF-Ts is more retarded on the polyacrylamide gel with respect to its *Escherichia coli* counterpart (compare lanes 6 and 12, respectively). This is probably because of the difference in length between *Escherichia coli* EF-Ts (283 aa residues) and *Geobacillus anatolicus* EF-Ts (294 aa residues + 6aa for His-tag). Bands appearing in lanes 2-5 and 8-11 showes EF-Tu·EF-Ts complexes. In these lanes and no free EF-Ts band is observed indicating that EF-Ts is fully active in EF-Tu binding (Figure 5.24).

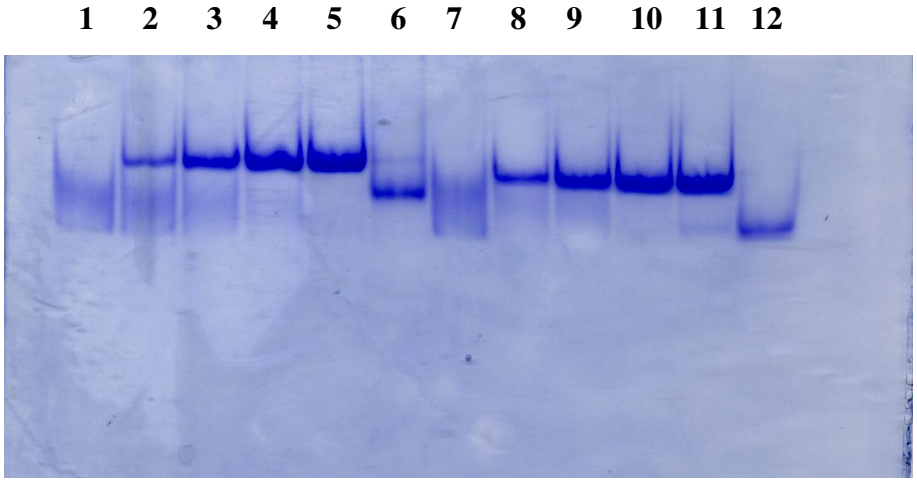

Figure 5.24. Results for binding assay for *Escherichia coli* EF-Tu and recombinant *Geobacillus anatolicus* EF-Ts. All lanes contain 300 pmol *Escherichia coli* EF-Tu except lane 6 and 12. Lanes 1 to 6, increasing recombinant *Geobacillus anatolicus* EF-Ts concentrations, (0 pmol, 100pmol, 200pmol, 300pmol, 390pmol, 390pmol), Lanes 7 to 12 increasing recombinant *Escherichia coli* EF-Ts concentrations, (0 pmol, 100pmol, 200pmol, 300pmol, 390pmol, 390pmol)

6. DISCUSSION

6.1. Determination of *Geobacillus anatolicus tsf* Gene Sequence

In order to determine the previously unknown gene sequence of *Geobacillus anatolicus* EF-Ts, Genbank database was searched and bacteria phylogenetically closely related to *Geobacillus anatolicus* were chosen for sequence determination study. Based on this data, degenerate primers were designed to amplify the partial *tsf* gene of *Geobacillus anatolicus*. The information about the genes that are upstream and downstream of *tsf* for *Geobacillus anotolicus* was not known. Therefore, the genes which are theoretically surrounding the *tsf* gene in the organisms that have full genome sequences were also aligned. Using the conserved regions in the upstream and downstream genes, the full *tsf* gene was amplified.

Full nucleotide sequence determination required six primers including those designed from the internal part of the *tsf* gene. Non-degenerate primers were designed after amplifying the partial *tsf* gene and its sequence determination. The full nucleotide sequence data was used to predict the amino acid sequence of the EF-Ts protein. The amino acid sequence data was consistent when compared with the other elongation factor Ts amino acid sequences.

The comparison between *Geobacillus anatolicus* and its closest relative *Geobacillus kaustophilus* EF-Ts amino acid sequences shows that 95.5 % of amino acids are conserved. This result verifies that the *Geobacillus anatolicus* is a new species based on 16S ribosomal DNA comparisons previously reported (Uysal *et al.,* 2001) (Figure 6.1). The nucleic acid and amino acid sequences of *Geobacillus anatolicus tsf* were the second G*eobacillus* species recorded for EF-Ts up to date.

Comparisons of various thermophilic organisms show that highly conserved amino acid region “TDFV” in EF-Ts sequences is also fully conserved in the *Geobacillus anatolicus* EF-Ts. The role of these residues is postulated to disrupt the Mg^{2+} binding capacity of EF-Tu and to lessen the binding affinity of EF-Tu to GDP so that the GDP

bound to EF-Tu after GTP hydrolysis can drop off to be replaced with GTP (Kawashima *et al.,* 1996, see also Zhang *et al.,* 1998). *Escherichia coli* EF-Tu·EF-Ts crystal structure is known (Kawashima *et al.,* 1996). There are 22 amino acid residues in EF-Ts in contact with EF-Tu according to this data (Kawashima *et al.,* 1996). The EF-Ts amino acid residues involved in interaction with EF-Tu according to this structural data are almost fully conserved in *Geobacillus anatolicus* EF-Ts (Figure 6.2). The alignment of *Escherichia coli* EF-Ts with *Geobacillus anatolicus* EF-Ts showed that only 4 of these amino acids (out of 22) are different (K-23 to Q-23, H-167 to D-179, M-235 to K-247 and M-278 to Q-290, between *Escherichia coli* EF-Ts and *Geobacillus anatolicus* EF-Ts, respectively). This corresponds to 82 % conservation. Whereas, in total amino acid sequence comparison between *Escherichia coli* EF-Ts and *Geobacillus anatolicus* EF-Ts, this conservation is 43 % (Figure 6.2).

```
                ....|....|  ....|....|  ....|....|  ....|....|  ....|....|
                    5           15          25          35          45
G kau           MAITAQMVKE LREKTGAGMM DCKKALTETN GDMEKAIDWL REKGIAKAAK
G ana           MAITAQMVKE LREKTGAGMM DCQKALTETN GDMEKAIDWL REKGIAKAAK
Consensus       MAITAQMVKE LREKTGAGMM DC KALTETN GDMEKAIDWL REKGIAKAAK

                ....|....|  ....|....|  ....|....|  ....|....|  ....|....|
                    55          65          75          85          95
G kau           KADRIAAEGM AYIAVEGNTA VILEVNSETD FVAKNEAFQT LVKELAAHLL
G ana           KRTASPAEGM TYIAVEATAA VILEVNSETD FVAKNEAFQT LVKELAAHLL
Consensus       K     AEGM  YIAVE   A VILEVNSETD FVAKNEAFQT LVKELAAHLL

                ....|....|  ....|....|  ....|....|  ....|....|  ....|....|
                   105         115         125         135         145
G kau           KQKPASLDEA LGQTMDNGST VQDYINEAIA KIGEKITLRR FAVVNKADGE
G ana           KQKPASLDEA LGQTMDNGST VQDYINEAIA KIGEKITLRR FAVVNKADGE
Consensus       KQKPASLDEA LGQTMDNGST VQDYINEAIA KIGEKITLRR FAVVNKADGE

                ....|....|  ....|....|  ....|....|  ....|....|  ....|....|
                   155         165         175         185         195
G kau           TFGAYLHMGG RIGVLTLLAG NASEDVAKDV AMHIAALHPK YVSRDDVPQE
G ana           TFGAYLHMGG RIGVLTLLAG NASEEIAKDV AMHIAALHPK YVSRDEVPQE
Consensus       TFGAYLHMGG RIGVLTLLAG NASE  AKDV AMHIAALHPK YVSRD VPQE

                ....|....|  ....|....|  ....|....|  ....|....|  ....|....|
                   205         215         225         235         245
G kau           EIAHEREVLK QQALNEGKPE KIVEKMVEGR LNKFYEDVCL LEQAFVKNPD
G ana           EIARREVLK QQALNEGKPE KIVEKMVEGR LNKFYEDVCL LEQAFVKNPD
Consensus       EIA EREVLK QQALNEGKPE KIVEKMVEGR LNKFYEDVCL LEQAFVKNPD

                ....|....|  ....|....|  ....|....|  ....|....|  ....
                   255         265         275         285
G kau           VTVRQYVESN GATVKQFIRY EVGEGLEKRQ DNFAEEVMSQ VRKQ
G ana           VTVRQYVESN GATVKQFIRY EVGEGLEKRQ DNFAEEVMSQ VRKQ
Consensus       VTVRQYVESN GATVKQFIRY EVGEGLEKRQ DNFAEEVMSQ VRKQ
```

Figure 6.1. Comparison of EF-Ts amino acid sequence between *Geobacillus anatolicus* and *Geobacillus kaustophilus* (non-conserved amino acids are shown in red)

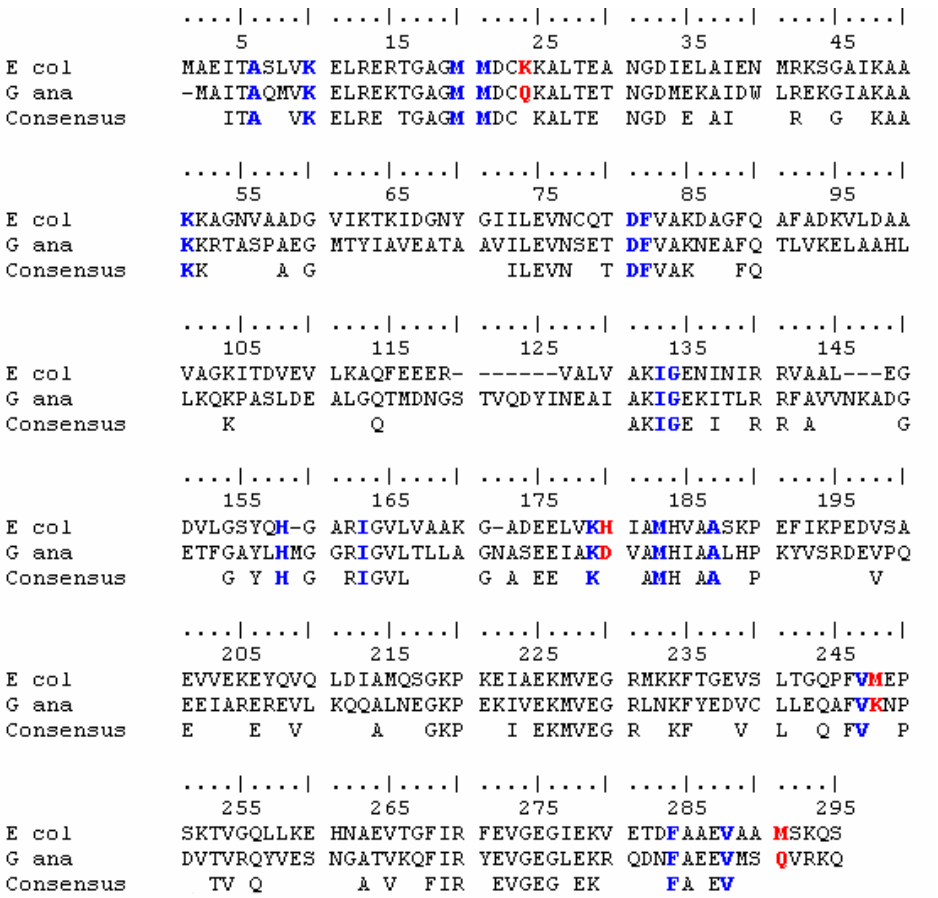

Figure 6.2. *Escherichia coli* EF-Ts and *Geobacillus anatolicus* EF-Ts
Residues that are in direct contact with EF-Tu according to the crystal structure data (Kawashima *et al.*, 1996) are colored (blue: conserved, red: non-conserved)

6.2. Cloning of the *tsf* Gene

The vector pCR® T7/NT-TOPO® was chosen for the cloning of the *tsf* gene. This vector contains multiple restriction enzyme sites which could be used for PCR cloning reactions. The vector also contains an enterokinase cleavage site. However, preliminary experiments showed that commercially available enterokinases cleave *Escherichia coli* EF-Ts even though this protein does not contain an enterokinase cleavage site probably due to unspecific cleavage activities of the recombinant commercial enzymes. For this reason, the enterokinase cleavage site was removed form the original vector and a new cloning vector was created with sticky ends of Nde I and Csp45 I. The primers were designed so that after double restriction digestions and purification, the PCR product was available to be inserted into the vector (See Figure 5.15 in Results).

6.3. Expression of the Recombinant Elongation Factor Ts

tsf gene is under the strong T7 RNA polymerase promoter in the expression vector, pT7D3Ts. However, wild type *Escherichia coli* do not express T7 RNA polymerase, so a special *Escherichia coli* strain, known as DE3, was used. DE3 is the designation for the lambda DE3 lysogen carrying the gene for T7 RNA polymerase under the control of the *lac*UV5 promoter. To overcome recombination based problems, a recombination deficient *Escherichia coli* strain JM109 (DE3) was chosen for the expression of the recombinant protein in this study. For the induction of the T7 RNA polymerase expression, IPTG was used after transforming *Escherichia coli* JM109 (DE3) cells with pT7D3Ts.

6.4. Purification of the Recombinant His-tagged Elongation Factor Ts

Metal-ion chelating chromotography was used for the purification of the recombinant His-tagged elongation factor Ts. The His-tag was artifically inserted into the protein at its C-terminal position by using the primer TsR2 (See TsR2, Figure 5.13 in Results). This insertion of His-tag gave the opportunity of purification of the recombinant protein by Ni^{2+}-affinity chromatography. Previous experiments showed that C-terminal His-tag on EF-Ts did not have negative effects on proteins' function (Jeppesen *et al*., 2005)

6.5. Binding assay of *Escherichia coli* EF-Tu with *Geobacillus anatolicus* EF-Ts

EF-Ts is the first of the elongation factors isolated from *Geobacillus anatolicus*. In order to study its binding to EF-Tu, no *Geobacillus anatolicus* EF-Tu was available. Therefore, the binding capacity of the purified EF-Ts was analyzed in its interaction with *Escherichia coli* EF-Tu. The binding assay results showed that *Geobacillus anatolicus* EF-Ts is fully active in binding to *Escherichia coli* EF-Tu. The titrations with EF-Ts at a constant *Escherichia coli* EF-Tu concentration indicated a 1:1 stoichiometry between EF-Tu and EF-Ts (Figure 5.24).

6.6. Codon Usage

Comparison with *Escherichia coli* codon usage data shows that the GC content of the *Geobacillus anatolicus* elongation factor Ts gene is slightly higher than *Escherichia coli*. In *Escherichia coli* total protein coding genes, the GC content was found to be 49,74 % and this ratio was found to be 51,4 % in *Geobacillus anatolicus* EF-Ts gene. The *Geobacillus anatolicus tsf* gene nucleotide GC content was significantly higher than the previously reported GC content for thermophiles, 47.9 % (Singer and Hickey, 2003). Codon preferences for the *tsf* gene from *Geobacillus anatolicus* do not vary significantly from the *Escherichia coli* codon preferences calculated for the whole genome (Table 6.1). *Geobacillus anatolicus* EF-Ts codon preferences are given in Table 6.2. In *Escherichia coli* genome the mostly used codon triplet is CUG which codes for Leu, whereas in Geobacillus anatolicus EF-Ts sequence GAA, which codes for Glu, is the mostly used codon.

UUU 22.6	UCU 11.0	UAU 18.5	UGU 5.4
UUC 15.6	UCC 9.3	UAC 12.0	UGC 6.0
UUA 15.1	UCA 10.1	UAA 2.0	UGA 1.0
UUG 12.9	UCG 8.5	UAG 0.3	UGG 13.8
CUU 12.7	CCU 8.0	CAU 12.5	CGU 18.8
CUC 10.1	CCC 5.6	CAC 8.8	CGC 18.2
CUA 4.6	CCA 8.7	CAA 14.4	CGA 4.1
CUG 45.6	CCG 19.1	CAG 28.0	CGG 6.6
AUU 29.6	ACU 11.1	AAU 23.1	AGU 10.8
AUC 22.5	ACC 21.2	AAC 21.1	AGC 14.9
AUA 8.5	ACA 10.9	AAA 35.6	AGA 4.6
AUG 25.8	ACG 13.8	AAG 13.2	AGG 2.6
GUU 20.1	GCU 17.5	GAU 32.9	GGU 24.8
GUC 14.0	GCC 23.7	GAC 18.8	GGC 25.4
GUA 12.0	GCA 21.7	GAA 37.9	GGA 10.8
GUG 23.2	GCG 27.6	GAG 18.9	GGG 11.6

Table 6.1. *Escherichia coli* codon preferences (per thousand)
(taken from http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=Escherichia+coli+%5Bgbbct%5D)

UUU	10.1	UCU	0	UAU	6.7	UGU	0
UUC	16.9	UCC	6.7	UAC	16.9	UGC	6.7
UUA	13.5	UCA	3.3	UAA	0	UGA	3.3
UUG	13.5	UCG	10.1	UAG	0	UGG	3.3
CUU	6.7	CCU	0	CAU	10.1	CGU	3.3
CUC	20.3	CCC	0	CAC	3.3	CGC	33.8
CUA	0	CCA	0	CAA	47.4	CGA	0
CUG	20.3	CCG	20.3	CAG	3.3	CGG	6.7
AUU	16.9	ACU	6.7	AAU	13.5	AGU	0
AUC	33.8	ACC	10.1	AAC	27.1	AGC	6.7
AUA	0	ACA	6.7	AAA	64.4	AGA	3.3
AUG	37.2	ACG	27.1	AAG	23.7	AGG	0
GUU	16.9	GCU	16.9	GAU	23.7	GGU	6.7
GUC	27.1	GCC	27.1	GAC	20.3	GGC	37.2
GUA	10.1	GCA	20.3	**GAA**	**88.1**	GGA	13.5
GUG	30.5	GCG	57.6	GAG	27.1	GGG	6.7

Table 6.2. *Geobacillus anatolicus* EF-Ts codon preference (per thousand)

The Table 6.3 shows mean values for mesophilic and thermophilic codon preferences for some codon triplets (Singer and Hickey, 2003). According to this data, *Geobacillus anatolicus* EF-Ts amino acid sequence shows low convergence to thermophilic codon preference mean values.

Codon	Mean (mesophiles)	Mean (thermophiles)	Mean difference
GGA	11.2	25.4	14.2
GAG	21.2	44.7	23.5
GCG	29.2	15.2	– 13.9
AGG	3.1	29.1	26.1
AGA	4.0	19.3	15.3
AAG	14.5	39.6	25.1
AAT	19.0	12.1	– 6.9
ATA	6.1	39.2	33.0
ATT	29.5	18.3	– 11.2
ACC	21.0	12.4	– 8.6
TAC	13.0	26.7	13.7
TTG	22.0	10.2	– 11.8
TTC	15.4	26.5	11.1
CGG	7.9	2.5	– 5.4
CGA	4.3	1.5	– 2.8
CGT	17.1	2.5	– 14.6
CGC	19.8	3.6	– 16.2
CAG	23.3	14.3	– 9.0
CAA	21.6	5.2	– 16.3
CAT	12.8	5.5	– 7.4
CTT	13.8	22.3	8.6
CTC	13.4	26.3	12.9

Table 6.3. Number of codons per thousand in mesophiles and thermophiles (Adapted from Singer and Hickey, 2003)

6.7. Conclusion and Future Perspective

The recombinant elongation factor Ts should further be investigated in its interactions with its natural substrate EF-Tu from a homologous organism. Currently *Geobacillus anatolicus* EF-Tu has been cloned (Barış Akalın, thesis, in progress). Kinetic properties of the *Geobacillus anatolicus* EF-Ts and EF-Tu now can be investigated.

Development of an *in vitro* thermophilic translation system would give an insight into how the speed and the accuracy of translation can be maintained at high temperatures. *Geobacillus anatolicus* was isolated at 98^{o}C and therefore is a valuable organism for this investigation.

The thermal stability of the recombinant elongation factor Ts should be studied *in vitro*. Elongation factor Ts from mesophilic bacteria is already known to be highly thermostable. How the catalytic activity of the thermophilic elongation factor Ts is affected at lower temperatures should be investigated.

There were only 5 crystal structures solved up to date for EF-Ts in complex with EF-Tu; *Homo sapiens* mitochondrial, *Bos taurus* mitochondrial, *Escheriscia coli*, *Thermus thermophilus* and *Saccharomyces cerevisiae*. The purified EF-Ts from *Geobacillus anatolicus* alone and also in complex with EF-Tu (after cloning and purification of EF-Tu) will be used for structural studies.

REFERENCES

An, G., D.S. Bendiak, L.A. Mamelak and J.D. Friesen, 1981, "Organization and nucleotide sequence of a new ribosomal operon in *Escherichia coli* containing the genes for ribosomal protein S2 and elongation factor Ts", *Nucleic Acids Research*, Vol. 9, pp. 4163-4172.

Andersen, G.R., P. Nissen and J. Nyborg, 2003, "Elongation factors in protein biosynthesis", *TRENDS in Biochemical Sciences,* Vol. 28, pp. 434-441.

Blumenthal, T., J. Douglass and D. Smith, 1977, "Conformational alteration of protein synthesis elongation factor EF-Tu by EF-Ts and kirromycin*", Natl. Acad. Sci. USA,* Vol. 74, pp. 3264-3267.

Bradford, M. M., 1976, "Bradford Assay", *Analytical Biochemistry*, Vol. 72, pp. 248-254.

Brock, S., K. Szkaradkiewicz and M. Sprinzl, 1998, "Initiation factors of protein biosynthesis in bacteria and their structural relationship to elongation and termination factors", *Molecular Microbiology*, Vol. 29, pp. 409-417

Clark, B.F.C., S. Thirup, M. Kjeldgaard and J. Nyborg, 1999, "Structural information for explaining the molecular mechanism of protein biosynthesis", *FEBS Letters*, Vol. 452, pp. 41-46.

Ganoza, M.C., M.C. Kiel and H. Aoki, 2002, "Evolutionary Conservation of Reactions in Translation", *Microbiology and Molecular Biology Reviews*, Vol. 66, pp. 460-485.

Hoogvliet, G., G.P. van Wesel and B. Kraal, 1999, "Evidence that a single EF-Ts suffices for the recycling of multiple and divergent EF-Tu species in *Streptomyces coelicolor* A3(2) and *Streptomyces ramocissimus*", *Microbiology*, Vol. 145, pp. 2293-2301.

Jeppesen, M.G., T. Navratil, L.L. Spremulli and J. Nyborg, 2005, "Crystal Structure of the Bovine Mitochondrial Elongation Factor Tu·Ts Complex", *The Journal of Biological Chemistry,* Vol. 280, pp. 5071-5081.

Kawashima, T., C. Berthet-Colominas, M. Wulff, S. Cusack and R. Leberman, 1996, "The structure of the *Escherichia coli* EFTu·EF-Ts complex at 2·5 Å resolution", *Nature*, Vol. 379, pp. 511-518.

Kisselev, L., M. Ehrenberg and L. Frolova, 2003, "Termination of translation: interplay of mRNA, rRNAs and release factors?", *The EMBO Journal*, Vol. 22, pp. 175-182.

Kumar, S. and R. Nussinov, 2001, "How do thermophilic proteins deal with heat?", *Cellular and Molecular Life Sciences*, Vol. 58, pp.1216-1233.

Low, S. C. and M. J. Berry**,** 1996, "Knowing when not to stop: selenocysteine incorporation in eukaryotes", *Trends in Biological Sciences*, Vol. 21**,** pp. 203–208.

Madigan, M.T. and J.M. Martinko, 2005, *Brock Biology of Microorganisms*, Eleventh Edition, Prentice Hall.

Marintchev, A. and G. Wagner, 2005, "Translation initiation: structures, mechanisms and evolution", *Quarterly Reviews of Biophysics,* Vol. 37, pp. 1-88.

Nilsson, J. and P. Nissen, 2005, "Elongation factors on the ribosome", *Current Opinion in Structural Biology*, Vol. 15, pp. 349-355.

Singer, A.C.G. and D.A. Hickey, 2003, "Thermophilic prokaryotes have characteristic patterns of codon usage, amino acid composition and nucleotide content", *Gene*, Vol. 317, pp. 39-47.

Tate, W. P. and C. M. Brown, 1992, "Translational termination: "stop" for protein synthesis or "pause" for regulation of gene expression", *Biochemistry*, Vol 31, pp. 2443–2450.

Uysal, H., S. Bakkal, D. Erturk and N. Bilgin, 2001, "*Geobacillus anatolicus* Strain SB 16S Ribosomal RNA Gene, Partial", (Unpublished) Bacterial GenBank Accession Number : AF411064.

Vetriani, C., D.L. Maeder, N. Tolliday, K.S.-P. Yip, T.J. Stillman, K.L. Britton, D.W. Rice, H.H. Klump and F.T. Robb, 1998, "Protein thermostability above 100°C: A key role for ionic interactions", *Proc. Natl. Acad. Sci. USA*, Vol. 95, pp. 12300-12305.

Vitagliano, L., M. Masullo, F. Sica, A. Zagari and V. Bocchini, 2001, "The crystal structure of Sulfolobus solfataricus elongation factor 1A in complex with GDP reveals novel features in nucleotide binding and exchange", *The EMBO Journal*, Vol. 20, pp. 5305–5311.

Zhang, Y., N.-J. Yu and L.L. Spremulli, 1998, "Mutational Analysis of the Roles of Residues in *Escherichia coli* Elongation Factor Ts in the Interaction with Elongation Factor Tu", *The Journal of Biological Chemistry*, Vol. 273, pp. 4556-4562.